STATISTICAL STABILITY OF EQUILIBRIUM STATES FOR INTERVAL MAPS

JORGE MILHAZES FREITAS AND MIKE TODD

ABSTRACT. We consider families of transitive multimodal interval maps with polynomial growth of the derivative along the critical orbits. For these maps Bruin and Todd have shown the existence and uniqueness of equilibrium states for the potential $\varphi_t : x \mapsto -t \log |Df(x)|$, for t close to 1. We show that these equilibrium states vary continuously in the weak* topology within such families. Moreover, in the case $t = 1$, when the equilibrium states are absolutely continuous with respect to Lebesgue, we show that the densities vary continuously within these families.

1. Introduction

One of the main goals in the study of Dynamical Systems is to understand how the behaviour changes when we perturb the underlying dynamics. In this paper, we examine the persistence of statistical properties of a multimodal interval map (I, f). In particular we are interested in the behaviour of the Cesàro means $\frac{1}{n}\sum_{k=0}^{n-1}\varphi\circ f^k(x)$ for a potential $\varphi : I \to \mathbb{R}$ for 'some' points x, as $n \to \infty$. If the system possesses an invariant *physical measure* μ, then part of this statistical information is described by μ since, by definition of physical measure, there is a positive Lebesgue measure set of points $x \in I$ such that

$$\overline{\varphi}(x) := \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \varphi \circ f^k(x) = \int \varphi \, d\mu.$$

If for nearby dynamics these measures are proven to be close, then the Cesàro means do not change much under small deterministic perturbations. This motivated Alves and Viana [AV] to propose the notion of *statistical stability*, which expresses the persistence of statistical properties in terms of the continuity, as a function of the map f, of the corresponding physical measures. A precise definition will be given in Section 1.1

However, the study of Cesàro means is not confined to the analysis of these measures. The study of other classes can be motivated through the encoding of these statistical

2000 *Mathematics Subject Classification.* 37D35, 37C75, 37E05, 37D25.

Key words and phrases. Equilibrium states, statistical stability, thermodynamic formalism, interval maps.

JMF is partially supported by POCI/MAT/61237/2004 and MT is supported by FCT grant SFRH/BPD/26521/2006. Both authors are supported by FCT through CMUP.

properties by 'multifractal decomposition', see [P] for a general introduction. Given $\alpha \in \mathbb{R}$, we define the sets

$$B_\varphi(\alpha) := \{x \in I : \overline{\varphi}(x) = \alpha\}, \ B'_\varphi := \{x \in I : \overline{\varphi}(x) \text{ does not exist}\}.$$

Then the multifractal decomposition in this case is

$$I = B'_\varphi \cup \left(\bigcup_\alpha B_\varphi(\alpha) \right).$$

Understanding the nature of this decomposition gives us information about the statistical properties of the system. This can be studied via 'equilibrium states'. See [PW] for a fuller account of these ideas, where the theory is applied to subshifts of finite type.

To define equilibrium states, given a potential $\varphi : I \to \mathbb{R}$, we define the *pressure* of φ to be

$$P(\varphi) := \sup \left\{ h_\mu + \int \varphi \, d\mu \right\},$$

where this supremum is taken over all invariant ergodic probability measures. Here h_μ denotes the metric entropy of the system (I, f, μ). Any such measure μ which 'achieves the pressure', i.e. $h_\mu + \int \varphi \, d\mu = P(\varphi)$, is called an *equilibrium state* for (I, f, φ).

In this paper for a given map f, we are interested in the equilibrium state μ_t of the 'natural' potential $\varphi_t : x \mapsto -t \log |Df(x)|$ for different values of t. For a multimodal map f and an f-invariant probability measure μ we denote the *Lyapunov exponent of* μ by $\lambda(\mu) := \int \log |Df| \, d\mu$. For any f in the class of multimodal maps $\mathcal{F}$ which we define below, Ledrappier [L] showed that for $t = 1$, there is an equilibrium state μ_1 with $\lambda(\mu_1) > 0$ if and only if μ_1 is absolutely continuous with respect to Lebesgue. We then refer to μ_1 as an acip. In this setting any acip is also a physical measure.

Using tools developed by Keller and Nowicki in [KN], Bruin and Keller [BK] further developed this theory, showing that for unimodal Collet-Eckmann maps there is an equilibrium state μ_t for φ_t for all t close to 1. This range of parameters was extended to all t in a neighbourhood of $[0, 1]$ for a special class of Collet-Eckmann maps by Pesin and Senti [PSe]. Bruin and the second author showed similar results for the non-Collet Eckmann multimodal case in [BT2].

The *Lyapunov exponent of a point* $x \in I$ is defined as $\overline{\varphi}_1(x)$, if this limit exists. So the set of points with the same Lyapunov exponent is $B_{\varphi_1}(\alpha)$. If f is transitive and there exists an acip then by the ergodic theorem $\mu_1(B_{\varphi_1}(\lambda(\mu_1)) = 1$. As shown in [T], under certain growth conditions on f, for a given value of α, close to $\lambda(\mu_1)$, there is an equilibrium state μ_t supported on $B_{\varphi_1}(\alpha)$ for some t close to 1. Therefore, to understand the statistics of the system with potential φ_1, it is useful to study the properties of the equilibrium states μ_t.

We would like to point out that for continuous potentials $\upsilon : I \to \mathbb{R}$ the theory of statistical stability has been studied in [Ar]. The strength of the current paper is that we deal with the more natural class of potentials φ_t, which are not continuous. By [BT2, Theorem 5] and [BT3, Theorem 6] under a growth condition on f, there

are equilibrium states for the potential $x \mapsto \varphi_t(x) + sv$ if v is Hölder continuous and (t, s) is close to $(1, 0)$ or to $(0, 1)$. For ease of exposition we will not consider this more general class of potentials here, but we note that the statistical stability for this more general class equilibrium states can be proved with only minor changes.

Our method provides a natural framework, via Gibbs states for inducing schemes, to study convergence of equilibrium states, broadening and simplifying previous approaches. Many of our proofs rely strongly on theory developed in [BLS] and [BT2]. Aside from the larger class of measures considered here, we are also able to study a class of maps with only polynomial growth along the critical orbit and hence, by [BLS], whose acips may only have polynomial decay of correlations. The maps considered here are multimodal, but we would also like to point out the theory presented in this paper extends to other cases such as Manneville-Pomeau maps, see Remark 6.2.

1.1. **Statement of results.** Here we establish our setting and make our statements more precise. Let Crit $=$ Crit(f) denote the set of critical points of f. We say that $c \in$ Crit is a *non-flat* critical point of f if there exists a diffeomorphism $g_c : \mathbb{R} \to \mathbb{R}$ with $g_c(0) = 0$ and $1 < \ell_c < \infty$ such that for x close to c, $f(x) = f(c) \pm |g_c(x - c)|^{\ell_c}$. The value of ℓ_c is known as the *critical order* of c. We define $\ell_{max}(f) := \max\{\ell_c : c \in \text{Crit}(f)\}$. Throughout $\mathcal{H}$ will be the collection of C^2 interval maps which have negative Schwarzian (that is, $1/\sqrt{|Df|}$ is convex away from critical points) and all critical points non-flat.

The *non-wandering set* is the set of points $x \in I$ such that for arbitrarily small neighbourhoods U of x there exists $n = n(U) \geqslant 1$ such that $f^n(U) \cap U \neq \emptyset$. This is the dynamically interesting set. As in [HR], for piecewise monotone C^2 maps this set splits into a possibly countable number of sets Ω_k on which f is topologically transitive. As in [MeS, Section III.4] (see also [BT2, Section 2.2]), these sets Ω_k can be Cantor sets or finite sets if the map is renormalisable: that is if there exists a cycle of intervals $J, f(J), \dots, f^p(J)$ so that $f^p(J) \subset J$, $f^p(\partial J) \subset \partial J$. The other possibility is that Ω_k is a cycle of intervals permuted by f. Our analysis extends to any such cycle of intervals, and indeed in that case we do not need to make any assumptions on critical points outside the cycle under consideration. However, for ease of exposition, we will assume that maps in $\mathcal{H}$ are non-renormalisable with only one transitive component Ω of the non-wandering set, a cycle of intervals. We also assume that for any $f \in \mathcal{H}$, $f^j(\text{Crit}) \cap f^k(\text{Crit}) \neq \emptyset$ implies $j = k$. For maps failing this assumption, either $f^k(\text{Crit}) \cap \text{Crit} \neq \emptyset$ for some $k \in \mathbb{N}$, in which case we could consider these relevant critical points in a block; or some critical point maps onto a repelling periodic cycle, which we exclude here for ease of exposition since our method is particularly tailored to case of more interesting maps where the critical orbits are infinite. It is also convenient to suppose that there are no points of inflection.

Let $\mathcal{H}_{r,\ell} \subset \mathcal{H}$ denote the set of maps $f \in \mathcal{H}$ with r critical points with $\ell_{max}(f) \leqslant \ell$. We will consider families of maps in $\mathcal{H}$ which satisfy the following conditions. The first one is the Collet-Eckmann condition: For any $r \in \mathbb{N}$, $\ell \in (1, \infty)$ and $C, \alpha > 0$,

the class $\mathcal{F}_e(r,\ell,C,\alpha)$ is the set

(1) $$f \in \mathcal{H}_{r,\ell} \text{ such that } |Df^n(f(c))| \geqslant Ce^{\alpha n} \text{ for all } c \in \text{Crit and } n \in \mathbb{N}.$$

Secondly we consider maps satisfying a polynomial growth condition: For any $r \in \mathbb{N}$, $\ell \in (1,\infty)$, $C > 0$, and any $\beta > 2\ell$, the class $\mathcal{F}_p(r,\ell,C,\beta)$ is the set

(2) $$f \in \mathcal{H}_{r,\ell} \text{ such that } |Df^n(f(c))| \geqslant Cn^{\beta} \text{ for all } c \in \text{Crit and } n \in \mathbb{N}.$$

We will take a map $f_0 \in \mathcal{F}$ where we suppose that either $\mathcal{F} = \mathcal{F}_e(r,\ell,C,\alpha)$ or $\mathcal{F} = \mathcal{F}_p(r,\ell,C,\beta)$, and consider the continuity properties of equilibrium states for maps in $\mathcal{F}$ at f_0.

We will consider equilibrium states for maps in these families. Suppose first that $\mathcal{F} = \mathcal{F}_e(r,\ell,C,\alpha)$. Then by [BT2, Theorem 2], there exists an open interval $U_{\mathcal{F}} \subset \mathbb{R}$ containing 1 and depending on α and r so that for $f \in \mathcal{F}$ and $t \in U_{\mathcal{F}}$ the potential $\varphi_{f,t} : x \mapsto -t\log|Df(x)|$ has a unique equilibrium state $\mu = \mu_f$. We note that by [MSz], the fact that there are r critical points gives a uniform upper bound $\log(r+1)$ on the topological entropy, which plays an important role in the computations which determine $U_{\mathcal{F}}$ in [BT2]. If instead we assume that $\mathcal{F} = \mathcal{F}_p(r,\ell,C,\beta)$ then by [BT2, Theorem 1] we have the same result but instead $U_{\mathcal{F}}$ is of the form $(t_{\mathcal{F}}, 1]$ where $t_{\mathcal{F}}$ depends on β, ℓ and r.

We choose our family $\mathcal{F}$, fix $t \in U_{\mathcal{F}}$ and denote $\varphi_{f,t}$ by φ_f. For every sequence $(f_n)_n$ of maps in $\mathcal{F}$ we let $\mu_{n,t} = \mu_{f_n,t}$ denote the corresponding equilibrium state for each n with respect to the potential φ_{f_n}. We fix $f_0 \in \mathcal{F}$ and say that $\mu_{0,t}$ is *statistically stable within the family* $\mathcal{F}$ if for any sequence $(f_n)_n$ of maps in $\mathcal{F}$ such that $\|f_n - f_0\|_{C^2} \to 0$ as $n \to \infty$, we have that $\mu_{0,t}$ is the weak* limit of $(\mu_{n,t})_n$.

Theorem A. *Let $\mathcal{F} \subset \mathcal{H}$ be a family satisfying* (1) *or* (2) *with potentials $\varphi_{f,t}$ as above. Then, for every fixed $t \in U_{\mathcal{F}}$ and $f \in \mathcal{F}$, the equilibrium state $\mu_{f,t}$ as above is statistically stable within the family $\mathcal{F}$.*

Although the definition of statistical stability involves convergence of measures in the weak* topology, when we are dealing with acips, it makes sense to consider a stronger type of stability due to Alves and Viana [AV]: for a fixed $f_0 \in \mathcal{F}$, we say that the acip μ_{f_0} is *strongly statistically stable* in the family $\mathcal{F}$ if for any sequence $(f_n)_n$ of maps in $\mathcal{F}$ such that $\|f_n - f_0\|_{C^2} \to 0$ as $n \to \infty$ we have

(3) $$\int \left| \frac{d\mu_{f_n}}{dm} - \frac{d\mu_{f_0}}{dm} \right| \, dm \xrightarrow[n\to\infty]{} 0,$$

where m denotes Lebesgue measure and μ_{f_n} and μ_{f_0} denote the acips for f_n and f_0 respectively. As a byproduct of the proof of Theorem A we also obtain:

Theorem B. *Let $\mathcal{F} \subset \mathcal{H}$ be a family satisfying* (1) *or* (2). *Then, for every $f \in \mathcal{F}$, the acip μ_f is strongly statistically stable.*

For uniformly hyperbolic maps, it is known that the measures do not merely vary continuously with the map, but actually vary differentiably in the sense of Whitney. For example, if $f_0 : M \to M$ is a C^3 Axiom A diffeomorphism of a manifold M with an unique physical measure μ_0, and the family $t \mapsto f_t$ is C^3, then the map

$t \mapsto \int \psi \, d\mu_t$ is differentiable at $t = 0$ for any real analytic observable $\psi : M \to \mathbb{R}$, see [Ru]. We would like to emphasise that the situation for non-uniformly hyperbolic maps is quite different. For example if $\mathcal{F}$ is the class of quadratic maps for which acips exist then it was shown in [Th] that these measures are not even continuous everywhere in this family, although as proved in [Ts] they are continuous on a positive Lebesgue measure set of parameters. It has been conjectured in [Ba1] that if $\mathcal{F}$ is the set of quadratic maps with some growth along the critical orbit then the acips should be at most Hölder continuous in this class, see also [Ba2]. For a positive result in that direction [RS] proved the Hölder continuity of the densities of the acips as in (3) for Misiurewicz parameters. Later, in [F], strong statistical stability was proved for Benedicks-Carleson quadratic maps, which are unimodal and satisfy condition (1). Hence, Theorem B provides a generalisation of this last result.

The following proposition shows that the pressure function is continuous in the family $\mathcal{F}$ with potentials $\varphi_{f,t}$. As with the proof of the main theorems, the proof uses the inducing structure and also the fact that we have some 'uniform decay rate' on these inducing schemes. We note that the continuity of the pressure for continuous potentials was related to statistical stability for non-uniformly expanding systems in [Ar].

Proposition 1.1. *Let $\mathcal{F}$ satisfy either* (1) *or* (2) *and let $t \in U_{\mathcal{F}}$. If $\{f_n\}_{n=0}^{\infty} \subset \mathcal{F}$ is such that $\|f_n - f_0\|_{C^2} \to 0$ as $n \to \infty$, then $P(\varphi_{f_n,t}) \to P(\varphi_{f_0,t})$ as $n \to \infty$.*

1.2. **Structure of the paper.** In Section 2, we build inducing schemes for each $f \in \mathcal{F}$ and show that the construction can be shadowed for nearby dynamics. Although other methods could be used to build the inducing schemes, we use Hofbauer towers. In Section 3, we introduce some thermodynamic formalism from [Sa4]. As proved in [BT2], this gives us the existence and uniqueness of equilibrium states for our inducing schemes for the relevant induced potentials. In particular, these equilibrium states satisfy the Gibbs property. In Section 4 we show that the weak* limit of Gibbs measures is also a Gibbs measure. We also prove Proposition 1.1, the continuity of the pressure over the family $\mathcal{F}$. In Section 5 this Gibbs measure is shown to be invariant. Finally, in Section 6 we show that the continuity of the measures survives the projection of the induced measures into the original equilibrium states, completing the proof of Theorem A. We finish that section by showing that the choice of inducing schemes and the uniformity properties of the family $\mathcal{F}$ proved along the way allow us to use the results of [AV] to obtain Theorem B.

We emphasise that the main new step in this paper is to use the fact, proved in [BT2], that the invariant measures on our inducing schemes are Gibbs states. This allows us to pass information from the limiting inducing scheme to other nearby inducing schemes. In this way we can avoid the techniques of [AV] which used convergence in the sense of (3). Those techniques can not be applied directly in this setting since, unless $\varphi = -\log |Df|$, we are not considering acips, and thus Lebesgue measure has no relevance.

In this paper we write $x = B^{\pm}y$ to mean $\frac{1}{B} \leqslant \frac{x}{y} \leqslant B$. For an interval J and a sequence of intervals $(J_n)_n$, we write $J_n \to J$ as $n \to \infty$ if the convergence is in the Hausdorff metric.

Acknowledgements: We would like to thank Paulo Varandas for helpful remarks, and Henk Bruin and Neil Dobbs for their comments on an early version of this paper. We would also like to thank anonymous referees for very useful comments.

2. Choice of inducing schemes

Given $f \in \mathcal{F}$, we say that (X, F, τ) is an *inducing scheme* for (I, f) if

- X is an interval containing a finite or countable collection of disjoint intervals X_i such that F maps each X_i diffeomorphically onto X, with bounded distortion (i.e. there exists $K > 0$ so that for all i and $x, y \in X_i$, $1/K \leqslant DF(x)/DF(y) \leqslant K$);
- $\tau|_{X_i} = \tau_i$ for some $\tau_i \in \mathbb{N}$ and $F|_{X_i} = f^{\tau_i}$.

The function $\tau : \cup_i X_i \to \mathbb{N}$ is called the *inducing time*. It may happen that $\tau(x)$ is the first return time of x to X, but that is certainly not the general case. For ease of notation, we will usually write $(X, F) = (X, F, \tau)$.

Given an inducing scheme (X, F, τ), we say that a measure μ_F is a *lift* of μ if for all μ-measurable subsets $A \subset I$,

$$\mu(A) = \frac{1}{\int_X \tau \, d\mu_F} \sum_i \sum_{k=0}^{\tau_i - 1} \mu_F(X_i \cap f^{-k}(A)). \tag{4}$$

Conversely, given a measure μ_F for (X, F), we say that μ_F *projects* to μ if (4) holds. We call a measure μ *compatible to* the inducing scheme (X, F, τ) if

- $\mu(X) > 0$ and $\mu\left(X \setminus \left\{x \in X : \tau(F^k(x)) \text{ is defined for all } k \geqslant 0\right\}\right) = 0$; and
- there exists a measure μ_F which projects to μ by (4); in particular $\int_X \tau \, d\mu_F < \infty$.

For the remainder of this paper we will denote the fixed map f in Theorems A and B by f_0 and take a sequence $(f_n)_n$ such that $f_n \in \mathcal{F}$ for all n and $\|f_n - f_0\|_{C^2} \to 0$.

The main purpose this section we use the theory of *Hofbauer towers* developed by Hofbauer and Keller [H, HK, K] to produce inducing schemes as described in [B]. We also show how the inducing schemes move with n. Note that we could also have used other methods to make these inducing schemes, see [BLS] for example.

We let $\mathcal{Q}_{n,k}$ be the natural partition into maximal closed intervals on which f_n^k is homeomorphic. We will denote members of $\mathcal{Q}_{n,k}$, which we refer to as *k-cylinders* by $\mathrm{X}_{n,k}$. Note that for $\mathrm{X}_{n,k}, \mathrm{X}'_{n,k} \in \mathcal{Q}_{n,k}$ with $\mathrm{X}_{n,k} \neq \mathrm{X}'_{n,k}$ then $\mathrm{X}_{n,k} \cap \mathrm{X}'_{n,k}$ consists of at most one point. For any $x \notin \cup_{j=0}^{k} f_n^{-j}(\mathrm{Crit}_n)$ there is a unique cylinder $\mathrm{X}_{n,j}$ containing x for $0 \leqslant j \leqslant k$. We denote this cylinder by $\mathrm{X}_{n,j}[x]$.

We next define the Hofbauer tower. We let

$$\hat{I}_n := \bigsqcup_{k\geqslant 0} \bigsqcup_{\mathrm{X}_{n,k}\in\mathcal{Q}_{n,k}} f_n^k(\mathrm{X}_{n,k})/\sim$$

where $f_n^k(\mathrm{X}_{n,k}) \sim f_n^{k'}(\mathrm{X}_{n,k'})$ if $f_n^k(\mathrm{X}_{n,k}) = f_n^{k'}(\mathrm{X}_{n,k'})$. Let $\mathcal{D}_n$ be the collection of domains of $\hat{I}_n$ and $\pi_n : \hat{I}_n \to I$ be the inclusion map. A point $\hat{x} \in \hat{I}_n$ can be represented by (x, D) where $\hat{x} \in D$ for $D \in \mathcal{D}_n$ and $x = \pi_n(\hat{x})$. In this case we can also write $D = D_{\hat{x}}$.

The map $\hat{f}_n : \hat{I} \to \hat{I}$ is defined as

$$\hat{f}_n(\hat{x}) = \hat{f}_n(x, D) = (f_n(x), D')$$

if there are cylinder sets $\mathrm{X}_{n,k} \supset \mathrm{X}_{n,k+1}$ such that $x \in f_n^k(\mathrm{X}_{n,k+1}) \subset f_n^k(\mathrm{X}_{k,n}) = D$ and $D' = f_n^{k+1}(\mathrm{X}_{n,k+1})$. In this case, we write $D \to D'$, giving $(\mathcal{D}_n, \to)$ the structure of a directed graph. We let $D_{n,0}$ denote the copy of $X_{n,0} = I$ in $\hat{I}_n$. For each $R \in \mathbb{N}$, let $\hat{I}_n^R$ be the compact part of the Hofbauer tower defined by

$$\hat{I}_n^R = \sqcup\{D \in \mathcal{D}_n : \text{ there exists a path } D_{n,0} \to \cdots \to D \text{ of length } r \leqslant R\}$$

The map π_n acts as a semiconjugacy between $\hat{f}_n$ and f_n: $\pi_n \circ \hat{f}_n = f_n \circ \pi_n$.

A subgraph $(\mathcal{E}, \to)$ of $(\mathcal{D}, \to)$ is called *closed* if $D \in \mathcal{E}$ and $D \to D'$ for some $D' \in \mathcal{D}$ implies that $D' \in \mathcal{E}$. It is *primitive* if for every pair $D, D' \in \mathcal{E}$, there is a path from D to D' within $\mathcal{E}$. Clearly any two distinct maximal primitive subgraphs are disjoint. We define $\mathcal{D}_{n,\mathcal{T}}$ to be the maximal primitive subgraph in $(\mathcal{D}_n, \to)$. We let $\hat{I}_{n,\mathcal{T}}$ be the union of all of these domains. This is the *transitive component of* $\hat{I}_n$. Since $f \in \mathcal{F}$ is transitive, there is a point $\hat{x} \in \hat{I}_{n,\mathcal{T}}$ so that $\overline{\bigcup_k \hat{f}_n^k(\hat{x})} = \hat{I}_{n,\mathcal{T}}$. The existence and uniqueness of this (maximal) component is implicit in works of Hofbauer and Raith [HR], see also [BT2, Lemma 1] for a self contained proof and references for the existence and uniqueness of this component and the existence of points with dense orbit.

We next explain how Hofbauer towers can be used to define inducing schemes. We use these schemes, rather than, for example, the very slightly different schemes in [BLS] since these were the schemes used in [BT2] and so we have good statistical information on them. This method of producing inducing schemes was first used in [B].

For an interval $A = (a, a + \gamma) \subset I$ and $\delta > 0$, we let $(1 + \delta)A$ denote the interval $(a - \delta\gamma, a + \gamma + \delta\gamma) \cap I$. Fixing $\delta > 0$ once and for all, for $A \subset I$ we let $A' = (1 + \delta)A$ and define

$$\check{A} = \check{A}_n(\delta) := \sqcup\{D \cap \pi_n^{-1}(A) : D \in \mathcal{D}_n, \pi_n(D) \supset A'\}. \tag{5}$$

Following the method of [B, Section 3], we pick some cylinder $\mathrm{X}_{n,k} \in \mathcal{Q}_{n,k}$ and consider the first return map $F_{\check{\mathrm{X}}_{n,k}} : \bigcup_j \hat{\mathrm{R}}^j \to \check{\mathrm{X}}_{n,k}$ where $\check{\mathrm{X}}_{n,k}$ is derived as in (5) and $F_{\check{\mathrm{X}}_{n,k}} = \hat{f}^{r_{\check{\mathrm{X}}_{n,k}}}$ for the return time $r_{\check{\mathrm{X}}_{n,k}}$ which is constant on each of the first return domains $\hat{\mathrm{R}}^j$. The fact, explained above, that each $\hat{f}_n$ is topologically

transitive on $\hat{I}_{n,\mathcal{T}}$ implies that $\overline{\bigcup_j \hat{\mathrm{R}}^j} = \check{\mathrm{X}}_{n,k}$. We next can define an inducing scheme $F_{\mathrm{X}_{n,k}} : \bigcup \mathrm{R}^j \to \mathrm{X}_{n,k}$ with inducing time $\tau_{\mathrm{X}_{n,k}}(x) = r_{\check{\mathrm{X}}_{n,k}}(\hat{x})$ for some $\hat{x} \in \check{\mathrm{X}}_{n,k}$ such that $\pi_n(\hat{x}) = x$. As shown in [B, Section 3], this number is the same for any such choice of $\hat{x}$. Here, after possibly relabelling, each $\hat{\mathrm{R}}^j$ is $\pi^{-1}(\mathrm{R}^j) \cap \check{\mathrm{X}}_{n,k}$. Moreover, $\tau_{\mathrm{X}_{n,k}}$ is constant $\tau^j_{\mathrm{X}_{n,k}}$ on each R^j. Let $(\mathrm{X}_{n,k})^\infty$ denote the set of points for which $\tau_{\mathrm{X}_{n,k}}(F^j_{\mathrm{X}^i_{n,k}}(x)) < \infty$ for all $j = 0, 1, \ldots$.

The main result of this section is the following proposition.

Proposition 2.1. *Let $\mathcal{F}$ be a fixed family $\mathcal{F} = \mathcal{F}_e(r, \ell, C, \alpha)$ or $\mathcal{F} = \mathcal{F}_p(r, \ell, C, \beta)$ satisfying* (1) *or* (2) *respectively. We let $(f_n)_n$ be any sequence such that $f_n \in \mathcal{F}$ for all n, and x be any point in $\Omega \setminus \left(\cup_{j\in\mathbb{Z}} f_0^j(\mathrm{Crit}_0)\right)$. For any k, if n is sufficiently large then there exists a sequence of inducing schemes $(\mathrm{X}_{n,k}[x], F_{\mathrm{X}_{n,k}[x]})$ as defined above such that $\mathrm{X}_{n,k}[x] \to \mathrm{X}_{0,k}[x]$ in the Hausdorff metric, and for any y in the interior of any R^j, $F_{\mathrm{X}_{n,k}[x]}(y) \to F_{\mathrm{X}_{0,k}[x]}(y)$ as $n \to \infty$.*

Note that the inducing schemes depend on $\delta > 0$, but any choice will work for all n.

Our inducing schemes are created as first return maps to sets $\check{\mathrm{X}}_{n,k}$ as defined above. Since the structure of these sets is determined by the structure of the Hofbauer tower, in order to prove the proposition we first have to show that the respective Hofbauer towers converge. Moreover, since the domains $\check{X}_{n,k}$ must be chosen inside $\hat{I}_{n,\mathcal{T}}$, we need to show that the sets $\hat{I}_{n,\mathcal{T}}$ converge. Without our assumptions on the uniform growth of all f_n, this may not be the case.

Lemma 2.2. *Let $\mathcal{F}$ be a fixed family $\mathcal{F} = \mathcal{F}_e(r, \ell, C, \alpha)$ or $\mathcal{F} = \mathcal{F}_p(r, \ell, C, \beta)$ satisfying* (1) *or* (2) *respectively. Then $\hat{I}_{n,\mathcal{T}} \to \hat{I}_{0,\mathcal{T}}$ in the sense that for any $R \in \mathbb{N}$, $\hat{I}^R_{n,\mathcal{T}} \to \hat{I}^R_{0,\mathcal{T}}$ in the Hausdorff metric.*

The proof of this lemma relies on the properties of measures on the Hofbauer tower, so before proving it, we show how to find representatives of these measures on the towers. Given $f \in \mathcal{F}$, we define $\iota := \pi|_{D_0}^{-1}$ where D_0 is the lowest level in $\hat{I}$, so $\iota : I \to D_0$ is an inclusion map. Given a probability measure m, let $\hat{m}^0 = m \circ \iota^{-1}$ be a probability measure on D_0. Let

$$\hat{m}^k := \frac{1}{k}\sum_{j=0}^{k-1} \hat{m}^0 \circ \hat{f}^{-j}.$$

We let $\hat{m}$ be a vague limit of this sequence. This is a generalisation of weak* limit for non-compact sets: for details see [K]. In general it is important to ensure that $\hat{m} \not\equiv 0$. It was proved in [BK, Theorem 3.6] that if m is an ergodic invariant measure with positive Lyapunov exponent then $\hat{m} \circ \pi^{-1} = m$. Note that for $f \in \mathcal{H}$, any $\hat{f}$-invariant probability measure $\hat{m}$, we must have $\hat{m}(\hat{I}_{\mathcal{T}}) = 1$.

Proof of Lemma 2.2. We first prove the following claim.

Claim 1. *$\hat{I}_n \to \hat{I}_0$ in the sense that for $R \in \mathbb{N}$, $\hat{I}^R_n \to \hat{I}^R_0$ in the Hausdorff metric.*

Proof. The idea of this proof is that since $\|f_n - f_0\|_{C^2} \to 0$, for any $k \in \mathbb{N}$ the set $\cup_{j=1}^k f_n^{-j}(\mathrm{Crit}_n)$ is topologically the same as the set $\cup_{j=1}^k f_0^{-j}(\mathrm{Crit}_0)$ for all n large. Observe that this is not necessarily true if we did not assume that critical orbits do not intersect, since otherwise there could be a point x such that $f_0^j(x) \in \mathrm{Crit}_0$ and $f_0^{j'}(x) \in \mathrm{Crit}_0$ for $j \neq j'$ and for each n two points $x_{n,1} \neq x_{n,2}$ such that $f_n^j(x_{n,1}) \in \mathrm{Crit}_n$ and $f_n^{j'}(x_{n,2}) \in \mathrm{Crit}_n$ and $x_{n,1}, x_{n,2} \to x$ as $n \to \infty$.

This means that for fixed k and for all large n, we can define an order preserving bijection on this set $h_{n,k} : \cup_{j=0}^k f_n^{-j}(\mathrm{Crit}_n) \to \cup_{j=0}^k f_0^{-j}(\mathrm{Crit}_0)$ so that for $x \in I$ such that $f_n^k(x) \in \mathrm{Crit}_n$, $f_0^k \circ h_{n,k}(x) \in \mathrm{Crit}_0$ and for all $0 \leqslant j \leqslant k$ we have $h_{n,k} \circ f_n^j(x) = f_0^j \circ h_{n,k}(x)$. Therefore, for any cylinder $\mathrm{X}_{0,k}^i \in \mathcal{Q}_{0,k}$ for large enough n, there is a corresponding cylinder $\mathrm{X}_{n,k}^i \in \mathcal{Q}_{n,k}$ so that $\mathrm{X}_{n,k}^i \to \mathrm{X}_{0,k}^i$. The existence of $h_{n,k}$ also implies that if, given $R \in \mathbb{N}$, for some $k, k' \leqslant R$, we have $f_0^k(\mathrm{X}_{0,k}^i) = f_0^{k'}(\mathrm{X}_{0,k'}^{i'})$ then for all large n, for the corresponding cylinders $\mathrm{X}_{n,k}^i$ and $\mathrm{X}_{n,k'}^{i'}$ we have $f_n^k(\mathrm{X}_{n,k}^i) = f_n^{k'}(\mathrm{X}_{n,k'}^{i'})$. Hence for fixed $R \in \mathbb{N}$, $\hat{I}_n^R$ and $\hat{I}_0^R$ are topologically the same for all large n. Since $\|f_n - f_0\|_{C^2} \to 0$ they also converge in the Hausdorff metric, completing the claim. □

Recall that our assumptions on the transitivity of maps in $\mathcal{H}$ and [BT2, Lemma 1] imply that there is a unique transitive component in the Hofbauer tower. The following claim will allow us to compare the transitive components of our Hofbauer towers.

Claim 2. *There is a domain $D_0^* \in D_{0,\mathcal{T}}$ so that the corresponding domains D_n^* are in $D_{n,\mathcal{T}}$.*

Assuming this claim, we use the fact that $(\mathcal{D}_{n,\mathcal{T}}, \to)$ is a closed subgraph, i.e. if $D \in \mathcal{D}_{n,\mathcal{T}}$ and there exists a path $D \to \cdots \to D'$ then $D' \in \mathcal{D}_{n,\mathcal{T}}$. Let D_0 be an arbitrary domain in $\mathcal{D}_{0,\mathcal{T}}$. By Claim 1, for large enough n, there exists a corresponding domain $D_n \in \mathcal{D}_n$. Since there must exist $\hat{x}_0 \in D_0^*$ with dense orbit in $\hat{I}_{0,\mathcal{T}}$, this point iterates into D_0 and back into D_0^*. Therefore, for large enough n there is a point $\hat{x}_n \in D_n^*$ which iterates into the corresponding domain $D_n \in \mathcal{D}_n$ and back out to D_n^*. Since $D_n^* \in \mathcal{D}_{n,\mathcal{T}}$, we must also have $D_n \in \mathcal{D}_{n,\mathcal{T}}$. Therefore, once the claim is proved, so is the lemma.

Proof of Claim 2. For $f \in \mathcal{F}$, and an open interval $U \subset I$, we say that x makes a *good entry to U at time k* if there exists an interval $U' \ni x$ so that $f^k : U' \to U$ is a homeomorphism. We first show that for the inducing domains Δ_n constructed as in [BLS], there exists $\theta > 0$ such that m-a.e. x makes a good entry to Δ_n under iteration by f_n with frequency at least θ. Here $\theta > 0$ is uniform in n and the measure m is Lebesgue. In fact we are really interested in good entries to a subset of Δ_n, but we do this first since it is simpler and introduces the ideas.

We let ν_n denote the acip for (I, f_n). In [BLS] inducing schemes $G_n : \bigcup \Delta_n^i \to \Delta_n$ are constructed for some Δ_n. Here $G_n = f^{r_n}$ for an inducing time r_n. We can take $\eta_n = \frac{|\Delta_n|}{2}$. This is uniformly bounded below, by some $\eta > 0$. To check this fact we

refer to [BLS, Lemma 4.2] where the sets Δ_n are constructed. Then observe that once a map f_0 is fixed, the construction of the corresponding Δ_0 involves a finite number of iterations and constants that can be taken uniformly within a neighbourhood of f_0. This means that one can mimic the construction for a neighbouring map f_n and hence obtain an interval uniformly close to the original Δ_0.

By the Ergodic Theorem, the frequency

$$\lim_{k\to\infty}\frac{1}{k}\#\left\{0\leqslant j<k:\exists U\ni x \text{ s.t. } f^j:U\to\Delta_n \text{ is a diffeomorphism}\right\}$$

for a ν_n-typical point x is bounded below by $1/\int r_n\,d\nu_{G_n}$ where ν_{G_n} is the measure for the inducing scheme as in (4). Since ν_n is equivalent to Lebesgue, we need only to show that $\int r_n\,d\nu_{G_n}$ is uniformly bounded above for all $f_n\in\mathcal{F}$, with n sufficiently large. This fact follows from Lemma 4.1 later in the paper.

As above, $\Delta_n\to\Delta_0$ in the Hausdorff metric. This means that there exists k and cylinders $\mathrm{X}_{n,k}\in\mathcal{Q}_{n,k}$ so that $\mathrm{X}_{n,k}\to\mathrm{X}_{0,k}$ and for all large n, $\mathrm{X}_{n,k}\Subset\Delta_n$. We set $D_n^*=f_n^k(\mathrm{X}_{n,k})$. A similar argument to the one above shows that m-a.e. x makes a good entry to $\mathrm{X}_{n,k}$ under iteration by f_n with frequency bounded below by $1/\int_{\mathrm{X}_{n,k}}r_n\,d\nu_{G_n}$. Since $\mathrm{X}_{n,k}$ converge to some $\mathrm{X}_{0,k}$ as in Claim 1, again Lemma 4.1 implies that for all large n, m-a.e. x makes a good entry to $\mathrm{X}_{n,k}$ under iteration by f_n with frequency bounded below by θ' where θ' is any value in $(0,1/\int_{\mathrm{X}_{0,k}}r_0\,d\nu_{G_0})$.

Given $f\in\mathcal{H}$, the way the Hofbauer tower is constructed using the cylinder structure means that if $U\subset I$ is an interval such that $f^k:U\to f^k(U)$ is a homeomorphism, then for $\hat U:=\iota(U)\subset D_0$, every iterate $\hat f^j(\hat U)$ is contained in a unique element of $\mathcal{D}$ for $0\leqslant j\leqslant k$. Moreover $\pi(\hat f^k(\hat U))=f^k(U)$. Therefore, if x makes a good entry to $\mathrm{X}_{n,k}$ at time j under iteration by f_n, then there exists an interval B in a unique domain of $\mathcal{D}_n$ so that $\pi_n(B)=\mathrm{X}_{n,k}$ and for $\hat x:=\iota_n(x)$, $\hat f_n^j(\hat x)\in B$. For any such a domain B, we must have $\hat f_n^k(B)=D_n^*$, and hence $\hat f_n^{j+k}(\hat x)\in D_n^*$.

Fix $\varepsilon>0$. The above argument means that the frequency of entries of a point $\iota_n(x)$ to D_n^* under iteration by $\hat f$ can be estimated in terms of good entries of x to $\mathrm{X}_{n,k}$ under iteration by f_n. Hence there exists $k_0=k_0(n,x,\varepsilon)\in\mathbb{N}$ so that $k\geqslant k_0$ implies

$$\frac{1}{k}\#\left\{0\leqslant j<k:\hat f^j(\hat x)\in\hat D_n^*\right\}>\frac{\theta'}{1+\varepsilon}.$$

Let N be so large that $m\{x\in I:k_0(n,x,\varepsilon)\leqslant N\}\geqslant 1-\varepsilon$. Then

$$\hat m^k(D_n^*)=\frac{1}{k}\sum_{j=0}^{k-1}\hat m^0\circ\hat f^{-j}(D_n^*)\geqslant\theta'\left(\frac{1-\varepsilon}{1+\varepsilon}\right)$$

for all $k\geqslant N$. Since $\varepsilon>0$ was arbitrary, we have $\hat m(D_n^*)\geqslant\theta'$ for all large n. Since $\hat m$ can only give mass to domains in transitive components, this implies that $D_n^*\subset\hat I_{n,\mathcal{T}}$ for all large n. □

□

Proof of Proposition 2.1. Lemma 1 of [BT2] implies that for any point $x \in \Omega \setminus \left(\cup_{k\in\mathbb{Z}} f_0^k(\mathrm{Crit}_0)\right)$ there exists $\hat{x} \in \hat{I}_{0,\mathcal{T}}$ so that $\pi_0(\hat{x}) = x$. Hence for large enough k, the cylinder $\mathrm{X}_{0,k}[x]$ will give rise to a set $\check{\mathrm{X}}_{0,k}[x] \subset \hat{I}_{0,\mathcal{T}}$ as in our construction which is non-empty. By Lemma 2.2, this will also be true of the corresponding cylinder for f_n for all n large enough. Moreover, that lemma implies that for any $R \in \mathbb{N}$, $\check{\mathrm{X}}_{n,k}[x] \cap \hat{I}_n^R \to \check{\mathrm{X}}_{0,k}[x] \cap \hat{I}_0^R$ as $n \to \infty$. Hence, the first return map by $\hat{f}_n$ to $\check{\mathrm{X}}_{n,k}[x]$ converges pointwise to that of $\hat{f}_0$ to $\check{\mathrm{X}}_{0,k}[x]$. Therefore for any $y \in \cup_j \overset{\circ}{R^j}$, $F_{\mathrm{X}_{n,k}[x]}(y) \to F_{\mathrm{X}_{0,k}[x]}(y)$ as $n \to \infty$, as required. $\square$

3. Equilibrium states for the induced maps

For a dynamical system $T : X \to X$ on a topological space and $\Phi : X \to \mathbb{R}$, we say that a measure m is Φ-*conformal* (and call Φ a potential) if $m(X) = 1$ and

$$m(T(A)) = \int_A e^{-\Phi(x)}\, dm(x)$$

whenever $T : A \to T(A)$ is one-to-one. In other words, $dm \circ T(x) = e^{-\Phi(x)} dm(x)$.

Assume that $\mathcal{S}_1 = \{\mathrm{C}_1^i\}_i$ is a countable Markov partition of X such that $T : \mathrm{C}_1^i \to X$ is injective for each $\mathrm{C}_1^i \in \mathcal{S}_1$. We denote $\mathcal{S}_k := \bigvee_{j=0}^{k-1} T^{-j}(\mathcal{S}_1)$, the k-join of the Markov partition $\mathcal{S}_1$, and suppose that the topology on X is generated by these sets. We define

$$V_k(\Phi) := \sup_{\mathrm{C}_k \in \mathcal{S}_k} \sup_{x,y \in \mathcal{S}_k} |\Phi(x) - \Phi(y)|, \tag{6}$$

We say that Φ has *summable variations* if $\sum_{k\geqslant 1} V_k(\Phi) < \infty$.

We define the transfer operator for a potential Φ with summable variations as

$$(\mathcal{L}_\Phi g)(x) := \sum_{T(y)=x} e^{\Phi(y)} g(y),$$

where g is in the Banach space of bounded continuous complex valued functions.

Suppose that (X, T) is topologically mixing and Φ is a potential with summable variations. For every $\mathrm{C}_1^i \in \mathcal{S}_1$ and $k \geqslant 1$ let

$$Z_k(\Phi, \mathrm{C}_1^i) := \sum_{T^k x = x} e^{\Phi_k(x)} 1_{\mathrm{C}_1^i}(x),$$

where $\Phi_k(x) = \sum_{j=0}^{k-1} \Phi \circ T^j(x)$. As in [Sa1], we define the *Gurevich pressure* of Φ as

$$P_G(\Phi) := \lim_{k\to\infty} \frac{1}{k} \log Z_k(\Phi, \mathrm{C}_1^i).$$

This limit exists since $\log Z_k(\Phi, \mathrm{C}_1^i)$ is almost superadditive:

$$\log Z_k(\Phi, \mathrm{C}_1^i) + \log Z_j(\Phi, \mathrm{C}_1^i) \leqslant \log Z_{k+j}(\Phi, \mathrm{C}_1^i) + \sum\nolimits_{n\geqslant 1} V_n(\Phi).$$

Therefore, $P_G(\Phi) = \sup_k \frac{1}{k} \log Z_k(\Phi, \mathrm{C}_1^i) > -\infty$. By the mixing condition, in [Sa1, Lemma 3], $P_G(\Phi)$ is independent of the choice of C_1^i. To simplify the notation, we

will often suppress the dependence of $Z_k(\Phi, \mathrm{C}_1^i)$ on C_1^i. Furthermore, if $\|\mathcal{L}_\Phi 1\|_\infty < \infty$ then $P_G(\Phi) < \infty$, see [Sa1, Lemma 2].

Assume now that $T : X \to X$ is the full shift. That is $T : \mathrm{C}_1^i \to X$ is bijective for all i.

We say that μ is a *Gibbs measure* if there exists $K < \infty$ such that for all $\mathrm{C}_k \in \mathcal{S}_k$,

$$\frac{1}{K} \leqslant \frac{\mu(\mathrm{C}_k)}{e^{\Phi_k(x) - kP_G(\Phi)}} \leqslant K$$

for any $x \in \mathrm{C}_k$. Here $\Phi_k(x) := \Phi(T^{k-1}(x)) + \cdots + \Phi(x)$.

Theorem 3.1 ([Sa4])**.** *If (X, T) is the full shift, $\Phi : X \to \mathbb{R}$ is a potential with $\sum_{k \geqslant 1} V_k(\Phi) < \infty$ and $P_G(\Phi) < \infty$ then*

(a) *There exists a unique Gibbs measure m_Φ, which is moreover $(\Phi - P_G(\Phi))$-conformal;*
(b) *There exists an invariant probability measure, which is also Gibbs, $\mu_\Phi = \rho_\Phi m_\Phi$ where ρ_Φ is unique so that $\mathcal{L}_\Phi \rho_\Phi = e^{P_G(\Phi)} \rho_\Phi$. Moreover, ρ_Φ is positive, continuous and has summable variations;*
(c) *If $h_{\mu_\Phi}(T) < \infty$ or $-\int \Phi d\mu_\Phi < \infty$ then μ_Φ is the unique equilibrium state (in particular, $P(\Phi) = h_{\mu_\Phi}(T) + \int_X \Phi \, d\mu_\Phi$);*
(d) *The Variational Principle holds, i.e., $P_G(\Phi) = P(\Phi)$ $(= h_{\mu_\Phi}(T) + \int_X \Phi \, d\mu_\Phi)$.*

Note that because μ_Φ is a Gibbs measure, $\mu_\Phi(\mathrm{C}_k^i) > 0$ for every cylinder set $\mathrm{C}_k^i \in \mathcal{S}_k$, $k \in \mathbb{N}$.

From Proposition 2.1, we have inducing schemes $(\mathrm{C}_{n,0}, F_n)$ for $\mathrm{C}_{n,0} = \mathrm{X}_{n,k}^i$ and $F_n = F_{\mathrm{X}_{n,k}^i}$. As in [BT2] we fix $t \in U_{\mathcal{F}}$ and let $\psi_n = \psi_{n,t} := \varphi_{f_n,t} - P(\varphi_{f_n,t})$. The corresponding induced potential is defined as $\Psi_n(x) = \Psi_{F_n}(x) := \psi_n \circ f^{\tau_n(x)}(x) + \cdots + \psi_n(x)$. These schemes can be coded symbolically by the full shift on countably many symbols. We denote a k-cylinder of F_n by $\mathrm{C}_{n,k}$, and the collection of these cylinders by $\mathcal{P}_{n,k}$. This plays the role of $\mathcal{S}_k$ in the discussion of the full shift above. We denote $\Psi_{n,k}(x) := \Psi_n(F_n^{k-1}(x)) + \cdots + \Psi_n(x)$. The variation $V_k(\Psi_n)$ is defined as in (6).

As shown in [PSe, BT2], Theorem 3.1 can then be used to produce equilibrium states for the systems $(\mathrm{C}_{n,0}, F_n, \Psi_n)$. Firstly, it can be shown, for example in [BT2, Lemma 7], that Ψ_n have summable variations. Next, in the proofs of Theorem 1 and 2 of [BT2] it was proved that $P_G(\Psi_n) = 0$ when f_n satisfies (2) and (1) respectively. Then using Theorem 3.1 we get conformal measures $m_{F_n} = m_{\Psi_n}$, densities $\rho_{F_n} = \rho_{\Psi_n}$, and equilibrium states $\mu_{F_n} = \rho_{F_n} m_{F_n}$ for $(\mathrm{C}_{n,0}, F_n, \Psi_n)$. These project to equilibrium states $\mu_n = \mu_{\psi_n}$ for (I, f_n, ψ_n). Note that an equilibrium state for ψ_n is also an equilibrium state for $\varphi_{f_n,t}$. Also note that these arguments imply that μ_{f_n} is compatible to each of the inducing schemes in Proposition 2.1.

We finish this section by proving a uniform bound on the variation of ρ_{F_n} which will be useful later.

Remark 3.2. *We define the distortion constants*

$$B_{n,k} := \exp\left(\sum\nolimits_{j\geqslant k+1} V_j(\Psi_n)\right).$$

By [BT2, Lemma 7] *the Koebe space given by the fact that our inducing schemes have diffeomorphic extensions to* $(1+\delta)C_{n,0}$ *implies that there exist* $0<\lambda(\delta,t)<1$ *and* $C(\delta)>0$ *so that* $V_k(\Psi_n)\leqslant C(\delta)\lambda(\delta,t)^k$*. Then there exist* $C'(\delta)>0$ *and* $\lambda'(\delta,t)$ *so that* $B_{n,k}=B_{n,k}(\delta,t)\leqslant C'(\delta)\exp\left(\sum\nolimits_{j\geqslant k+1}\lambda'(\delta,t)^k\right)$*. Therefore we can choose* $B_{n,k}$ *to be independent of* n*. We denote this bound by* B_k*.*

Following the proof of [Sa4, Theorem 1], any constant H_n with $H_n\geqslant(\sup\rho_{F_n})^2$ where ρ_{F_n} is as in Theorem 3.1(b) has the following property. For any $\mathrm{C}_{n,k}\in\mathcal{P}_{n,k}$,

$$\frac{1}{H_nB_0}\leqslant\frac{\mu_{F_n}(\mathrm{C}_{n,k})}{e^{\Psi_{n,k}(x)-kP_G(\Psi_n)}}\leqslant H_nB_0$$

for any $x\in\mathrm{C}_{n,k}$. As noted above, by [BT2], $P_G(\Psi_n)=0$. We are allowed to take a uniform distortion constant B_0 for all of our maps F_n by our choice of $\mathrm{C}_{n,0}$. It is important here to replace H_n with a uniform constant H. We consider how H_n was obtained. For the following lemma and its proof we fix $f=f_n$, so dropping any extra notation. Note that the bound used in [Sa4, p1754] is not sufficient for us since it depends on the measure of a cylinder, which can be different for different n.

Lemma 3.3. $V_0(\log\rho_\Psi)\leqslant 2\log B_0$ *and the Gibbs constant can be chosen to be* $H_f=(B_0)^4$*.*

Proof. According to [Sa2, (3.12)], see also the argument of [Sa1, Lemma 6], $V_1(\log\rho_\Psi)<\log B_1$. We use this to show $V_0(\log\rho_\Psi)$ is uniformly bounded above. Take $x_1,x_2\in\mathrm{C}_0$. Let $y_{1,i},y_{2,i}$ be the unique points in C_1^i such that $F(y_{1,i})=x_1$ and $F(y_{2,i})=x_2$. Then since $\mathcal{L}_\Psi\rho_\Psi=\rho_\Psi$,

$$\left|\frac{\rho_\Psi(x_1)}{\rho_\Psi(x_2)}\right|=\left|\frac{\sum_{Fy_1=x_1}e^{\Psi(y_1)}\rho_\Psi(y_1)}{\sum_{Fy_2=x_2}e^{\Psi(y_2)}\rho_\Psi(y_2)}\right|\leqslant\left|\frac{\sum_i\sup_{x\in\mathrm{C}_1^i}e^{\Psi(x)}\rho_\Psi(x)}{\sum_i\inf_{y\in\mathrm{C}_1^i}e^{\Psi(y)}\rho_\Psi(y)}\right|\leqslant B_0B_1.$$

Therefore the first part of the lemma is finished. There must exist $x_1,x_2\in\mathrm{C}_0$ with $\rho_\Psi(x_1)\leqslant1$ and $\rho_\Psi(x_2)\geqslant1$: otherwise in the first case $\mu_F(\mathrm{C}_0)>1$, and in the second case $\mu_F(\mathrm{C}_0)<1$. So setting $H_f:=(B_0)^4$ we have $H_f\geqslant(\sup\rho_\Psi)^2$, so we are finished. □

For use later, we let $H_{\mathcal{F}}:=(B_0)^4$.

4. Gibbs property for the weak* limit of Gibbs measures

Later in this section we will fix some inducing schemes (X_n,F_n) as in Proposition 2.1 with induced measures μ_{F_n}. By passing to a subsequence if necessary, we can define μ_{F_∞}, a weak* limit of $(\mu_{F_n})_n$. From the previous section and a uniqueness argument from [MU], we know that if we prove that μ_{F_∞} satisfies the Gibbs property and is invariant, then $\mu_{F_\infty}=\mu_{F_0}$. This section is devoted to proving that m_{F_∞} has the

Gibbs property which will allow us to conclude that μ_{F_∞} has the Gibbs property also. The proof of the following lemma relies heavily on [BLS] and [BT2]. In the proof we outline the main ideas used from those papers, in particular the origin of the important constants used.

Lemma 4.1. *For a given family $\mathcal{F} = \mathcal{F}_e(r, \ell, C, \alpha)$ or $\mathcal{F} = \mathcal{F}_p(r, \ell, C, \beta)$ satisfying* (1) *or* (2) *respectively, and every f in a neighbourhood of any $f_0 \in \mathcal{F}$ fixed, there exist $C' > 0$, $\alpha' > 0$ or $\beta' > 0$ and an inducing scheme (X, F) as in Proposition 2.1, with inducing time τ, so that for all $N \geqslant 1$,*

$$\mu_F\{\tau > N\} \leqslant C' e^{-\alpha' N} \text{ or } \mu_F\{\tau > N\} \leqslant C' N^{-\beta'} \text{ respectively .}$$

Proof. In the proof of Theorem 1 of [BT2], the correspondence between our inducing scheme and the one considered in [BLS] is given, which allows to conclude that the estimates for the tail of our inducing scheme in the case the potential is $\psi_{n,1}$, i.e. for the acip, are given by the ones in [BLS]. In the proofs of Theorems 1 and 2 of [BT2] it is then shown how estimates for the tails for the potentials $\psi_{n,t}$ can be easily related. The main result in [BLS] is proven by establishing that the growth rate of the derivative along the critical orbits determines the rate at which the Lebesgue measure of the tail of the inducing scheme falls off with time. Hence, roughly speaking, the estimates for the tail obtained in [BLS] depend essentially on the parameters that define the family $\mathcal{F}$. We will give some more insight on the construction of the inducing schemes of [BLS] so that the role of the constants involved becomes clearer. In what follows we will use the notation of [BLS] although it may differ from the one we use in the rest of the paper.

Bruin, Luzzatto and van Strien build a Markov map $f^R : \Omega_0 \to \Omega_0$ on a small neighbourhood of one of the critical points. The key idea is that outside a neighbourhood Δ of the critical points we have hyperbolic behaviour which leads to the exponential growth of the derivative. On the other hand, when we enter Δ, which happens frequently, we have a serious setback on hyperbolicity since the derivative takes values very close to 0. However, because inside Δ points are very close to a critical point they initiate a bind to it and regain hyperbolicity on account of the derivative growth experienced by the critical orbits which we have assumed. Once this binding ends, the losses are fully recovered and the derivative then grows exponentially fast until we enter Δ again and the cycle repeats itself. This means that the amount of time spent before an interval from Ω_0 becomes large enough to cover the whole Ω_0, reflects the growth rate of the derivative along critical orbits. The construction of the full return map is made in three major steps.

The first step is carried out in [BLS, Section 2], where the binding argument is described which allows inducing to small scales. In this section, the following constants are introduced: κ the *bounded backward contraction constant* (see [BT2, Lemma 9]), which depends only on the parameters that define the family $\mathcal{F}$ and on the number of critical points; K_0 a Koebe distortion constant that turns out to be ≤ 16; δ which establishes the size of the critical neighbourhood Δ and depends on κ, K_0, and on the parameters that define the family $\mathcal{F}$. This means that these constants can be chosen uniformly within the family. Some expansion estimates are also derived in this section. In [BLS, Lemma 2.4] the constants C_δ and λ_δ, that essentially give the

hyperbolicity outside the critical region, can be chosen uniformly inside a neighbourhood of a fixed $f_0 \in \mathcal{F}$. The crucial estimate that gives the recovery of hyperbolicity after the binding period is established in [BLS, Lemma 2.5]. These estimates depend only on the parameters that define the family $\mathcal{F}$.

The second step, the most influential step for the tail rate estimates, is done in [BLS, Section 3]. It consists in inducing to large scales which is stated in Proposition 3.1. Essentially, it is proved that there exists $\delta' > 0$ such that for all $\delta'' > 0$, any interval J of length at least $\delta'' > 0$, can be partitioned in such a way that for every element ω of the partition, there is a stopping time $\hat{p}(\omega)$ such that $f^{\hat{p}(\omega)}$ sends ω diffeomorphically onto an interval of length at least δ'. Moreover, the tail of this stopping time function, $m(x \in J : \hat{p}(x) > n)/m(J)$ decays accordingly to the derivative expansion on the critical orbits. The constant δ' is given by the contraction principle and can be taken uniformly inside a neighbourhood of any fixed $f_0 \in \mathcal{F}$. During this section, combinatorial estimates are obtained and several constants that can be chosen uniformly are introduced. The crucial observation is that the constants obtained for the estimates for $m(x \in J : \hat{p}(x) > n)/m(J)$ depend on the parameters that define the family $\mathcal{F}$, the constants fixed in the previous step and on both δ' and δ''. This means that the estimates on the tail of the time spent to reach large scale are uniform on a neighbourhood of any $f_0 \in \mathcal{F}$.

The third step, in [BLS, Section 4], gives the final construction of the full return map. It starts with [BLS, Lemma 4.2] that fixes the base Ω_0 for the inducing scheme that, as we have discussed in the proof of Claim 2, can be chosen uniformly inside a neighbourhood of any fixed $f_0 \in \mathcal{F}$. Then it is proved that once an interval achieves large scale, a fixed proportion of it will make a full return to the base Ω_0 in a finite number of iterates, which is a property that persists under small perturbations of f_0. This means that once an interval achieves large scale, it will make a full return exponentially fast. This implies that the estimates on the tail of the full return map are essentially the ones obtained in the second step for the time it takes to reach large scale. The constant δ'' is fixed and its value depends on δ' and on the size of Ω_0. All the other constants that appear turn out to depend on the parameters that define the family $\mathcal{F}$ and on previous constants, which means that they can all be chosen uniformly inside a neighbourhood of any fixed $f_0 \in \mathcal{F}$.

Note that in [BT2, Lemma 9] the condition that all the critical points had to have the same critical order, which had been required in [BLS], was removed. $\square$

As a consequence of this lemma, for a given $f_0 \in \mathcal{F}$ we can choose $\kappa = \kappa_{f_0} : \mathbb{N} \to [0,1]$ to be the function so that $\mu_{F_n}\{\tau_n > s_0\} \leqslant \kappa(s_0)$ for n large enough, and $\kappa(s_0) \to 0$ as $s_0 \to \infty$.

We next make conditions on our inducing schemes, so that only some of those in Proposition 2.1 will be appropriate choices. We select our inducing schemes so that the boundary of any 1-cylinder is accumulated by other 1-cylinders. In particular so that the boundary of a 1-cylinder with a small inducing time is accumulated by 1-cylinders with larger and larger inducing times.

Since $\mathrm{X}_{n,k} \in \mathcal{P}_{n,k}$ are cylinder sets, $f_n^j(\partial \mathrm{X}_{n,k}) \cap \overset{\circ}{\mathrm{X}}_{n,k} = \emptyset$ for all $1 \leqslant j \leqslant k$. However we can choose $\mathrm{X}_{0,k}^i \in \mathcal{P}_{0,k}$ so that $f_0^j(\partial \mathrm{X}_{0,k}^i) \cap \partial \mathrm{X}_{0,k}^i = \emptyset$ for all $1 \leqslant j \leqslant k$ also. Then as in Lemma 2.2, for each n large enough, there are corresponding cylinders $\mathrm{X}_{n,k}^i$ with $f_n^j(\partial \mathrm{X}_{n,k}^i) \cap \partial \mathrm{X}_{n,k}^i = \emptyset$ for all $1 \leqslant j \leqslant k$ also. It is easy to show that this property can be satisfied for our class of maps. We denote $\mathrm{C}_{n,0}$ to be the cylinder $\mathrm{X}_{n,k}^i$, which is fixed for the rest of this paper. The maps $F_n = F_{\mathrm{X}_{n,k}^i}$ are defined in Proposition 2.1. Recall that we set $\mathcal{P}_{n,0} := \{\mathrm{C}_{n,0}\}$, and define $\mathcal{P}_{n,k}$ to be the set of k-cylinders for the inducing scheme F_n. This construction means that $\mathrm{C}_{n,1}^{i_1} \cap \mathrm{C}_{n,1}^{i_2} = \emptyset$ for all $i_1 \neq i_2$ for all large n. We exploit this property in Remark 4.2. We may assume that this property actually holds for all n.

Let $\tau_{n,k}^i$ be the kth inducing time on a cylinder $\mathrm{C}_{n,k}^i$, i.e. $f^{\tau_{n,k}^i}(\mathrm{C}_{n,k}^i) = \mathrm{C}_{n,0}$. For brevity we will write $\tau_{n,1}^i = \tau_n^i$. For use later, note that our potentials $\Psi_{n,k}$ can be written as $\Psi_{n,k}(x) = \Phi_{n,k}(x) - P(\varphi_{f_n,t})\tau_{n,k}(x)$ where $\Phi_{n,k}(x) := -\log|DF_n^k(x)|$.

Any element of $\mathrm{C}_{0,k}^i \in \mathcal{P}_{0,k}$ is of the form $\mathrm{C}_{0,k}^i = [a_0, a_1]$ where there is some $p_0, p_1 \in \{0, 1, \ldots\}$ such that $f_0^{p_i}(a_i) \in \mathrm{Crit}_0$ for $i = 0, 1$. As in the proof Lemma 2.2, for n large enough, depending on $p := \max\{p_1, p_2\}$, there exists an order preserving bijection $h_{n,p} : \cup_{j=0}^p f_n^{-j}(\mathrm{Crit}_n) \to \cup_{j=0}^p f_0^{-j}(\mathrm{Crit}_0)$. Hence there are corresponding points $a_i^n := h_{n,p}^{-1}(a_i)$, $i = 0, 1$. By the proof of Proposition 2.1, there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, $[a_0^n, a_1^n]$ is a member $\mathcal{P}_{n,k}$, which we can label $\mathrm{C}_{n,k}^i$. We say that for $n \geqslant N$, $\mathrm{C}_{0,k}^i$ is *matched*; or similarly that $\mathrm{C}_{n,k}^i$ is *matched*. In this case, $\mathrm{C}_{n,k}^i \to \mathrm{C}_{0,k}^i$ as $n \to \infty$.

Remark 4.2. *Given $i \geqslant 1$, for all $M \geqslant 1$ there exists $\eta > 0$ and $N \geqslant 1$ so that for all $n \geqslant N$, $(1+\eta)\mathrm{C}_{n,1}^i \setminus \mathrm{C}_{n,1}^i$ only intersects 1-cylinders with $\tau_n > M$. To show this, we start by choosing N so large that $\{\mathrm{C}_{n,1}^j : \tau_n^j \leqslant M\}$ are matched for all $n \geqslant N$. Now let $\eta := \frac{1}{2}\min_{j \neq i,\ \tau_0^j \leqslant M} d(\mathrm{C}_{0,1}^i, \mathrm{C}_{0,1}^j)$. By the setup, $\eta > 0$. Now we may increase N so that $n \geqslant N$ implies $\mathrm{C}_{n,1}^j \cap \left(1 + \frac{\eta}{2}\right)\mathrm{C}_{0,1}^j = \mathrm{C}_{n,1}^j$ for all j with $\tau_0^j \leqslant M$. This means that η has the property required.*

Lemma 4.3. *For all $\varepsilon > 0$ there exists $i_0 \geqslant 1$ and $N \geqslant 1$ such that $\mathrm{C}_{0,1}^i$ is matched for all $1 \leqslant i \leqslant i_0$ for all $n \geqslant N$, and furthermore $n \geqslant N$ implies $\mu_{F_n}\left(\bigcup_{i > i_0} \mathrm{C}_{n,1}^i\right) < \varepsilon$.*

Proof. Let s_0 be so that $\kappa(s_0) < \varepsilon$. So s_0 depends only on ε and f_0 as in Lemma 4.1. We choose i_0 so that $\tau_0^i > s_0$ for all $i > i_0$. Similarly to Remark 4.2, we can choose N so large that $\mathrm{C}_{n,1}^i$ are matched for all $1 \leqslant i \leqslant i_0$ and that $\tau_n^i > s_0$ for all $i > i_0$ and all $n \geqslant N$. It then follows that $\mu_{F_n}\left(\bigcup_{i > i_0} \mathrm{C}_{n,1}^i\right) < \varepsilon$ as required. □

In the following lemmas we repeatedly use the conformal property of m_{F_n} for $n = 0, 1, 2 \ldots$. This allows us to compare behaviour at small scales with that at large scale.

Lemma 4.4. *For all $\varepsilon > 0$ for all $i_0 \geqslant 1$ there exists $\eta > 0$, such that for all $k \geqslant 1$, any $\mathrm{C}_{0,k}^j \in \mathcal{P}_{0,k}$ with $F_0^{k-1}(\mathrm{C}_{0,k}^j) = \mathrm{C}_{0,1}^i$ and $1 \leqslant i \leqslant i_0$ has*

$$\frac{m_{F_0}\left((1+\eta)\mathrm{C}_{0,k}^j\right)}{m_{F_0}(\mathrm{C}_{0,k}^j)} \leqslant B_0\left(1+\frac{\varepsilon}{4}\right) \text{ and } \frac{m_{F_0}\left(\left(\frac{1}{1+\eta}\right)\mathrm{C}_{0,k}^j\right)}{m_{F_0}(\mathrm{C}_{0,k}^j)} \geqslant \frac{1}{B_0\left(1+\frac{\varepsilon}{4}\right)}.$$

Proof. Let $s_1 \geqslant 1$ be such that

$$\kappa(s_1) \leqslant \frac{\varepsilon}{4}\left(\min_{1\leqslant i\leqslant i_0} m_{F_0}(\mathrm{C}_{0,1}^i)\right).$$

For the upper bound, let $\eta' > 0$ be such that the set $\bigcup_{1\leqslant i\leqslant i_0}(1+B_0\eta')\mathrm{C}_{0,1}^i \setminus \mathrm{C}_{0,1}^i$ contains only cylinders $\mathrm{C}_{0,1}^i$ with $\tau_0^i \geqslant s_1$ and is contained in $\mathrm{C}_{0,0}$. Then $m_{F_0}((1+B_0\eta')\mathrm{C}_{0,1}^i) \leqslant \left(1+\frac{\varepsilon}{4}\right) m_{F_0}(\mathrm{C}_{0,1}^i)$ for $1 \leqslant i \leqslant i_0$.

For $k > 1$, we use distortion and conformality to reduce the problem to the 1-cylinders' case just considered. Assume for $k \geqslant 1$ that $F_0^{k-1}(\mathrm{C}_{0,k}^j) = \mathrm{C}_{0,1}^i$. Since $(1+\eta')\,\mathrm{C}_{0,k}^j$ is in the same $k-1$-cylinder as $\mathrm{C}_{0,k}^j$, bounded distortion implies that

$$F_0^{k-1}\left((1+\eta')\,\mathrm{C}_{0,k}^j\right) \subset (1+B_0\eta')\mathrm{C}_{0,1}^i.$$

Using the conformal property of m_{F_0} and bounded distortion we have

$$\frac{m_{F_0}\left((1+B_0\eta')\,\mathrm{C}_{0,1}^i\right)}{m_{F_0}\left(\mathrm{C}_{0,1}^i\right)} \geqslant \frac{\int_{(1+\eta')\mathrm{C}_{0,k}^j} e^{-\Psi_{0,k-1}}\,dm_{F_0}}{\int_{\mathrm{C}_{0,k}^j} e^{-\Psi_{0,k-1}}\,dm_{F_0}} \geqslant \frac{1}{B_0}\left(\frac{m_{F_0}\left((1+\eta')\,\mathrm{C}_{0,k}^j\right)}{m_{F_0}(\mathrm{C}_{0,k}^j)}\right).$$

Hence, by the choice of η' above, we have

$$\frac{m_{F_0}\left((1+\eta')\,\mathrm{C}_{0,k}^j\right)}{m_{F_0}(\mathrm{C}_{0,k}^j)} \leqslant B_0\left(1+\frac{\varepsilon}{4}\right).$$

For the lower bound, let $s_2 \geqslant 1$ be such that $\kappa(s_2) < \frac{\varepsilon}{8}$. Then we choose $0 < \eta \leqslant \eta'$ so that the set $\mathrm{C}_{0,0} \setminus \frac{\mathrm{C}_{0,0}}{1+B_0\eta}$ only contains 1-cylinders $\mathrm{C}_{0,1}^i$ with $\tau_0^i \geqslant s_2$. This implies $m_{F_0}\left(\left(\frac{1}{1+B_0\eta}\right)\mathrm{C}_{0,0}\right) \geqslant 1-\frac{\varepsilon}{8} > \frac{1}{\left(1+\frac{\varepsilon}{4}\right)}$.

For $k > 1$ we use the a distortion argument similar to the one above. Bounded distortion implies that

$$F_0^k\left(\left(\frac{1}{1+\eta}\right)\mathrm{C}_{0,k}^j\right) \supset \left(\frac{1}{1+B_0\eta}\right)\mathrm{C}_{0,0}.$$

Using the conformal property of m_{F_0} and bounded distortion we have

$$\frac{m_{F_0}\left(\left(\frac{1}{1+B_0\eta}\right)\mathrm{C}_{0,0}\right)}{m_{F_0}\left(\mathrm{C}_{0,0}\right)} \leqslant \frac{\int_{\left(\frac{1}{1+\eta}\right)\mathrm{C}_{0,k}^j} e^{-\Psi_{0,k}}\,dm_{F_0}}{\int_{\mathrm{C}_{0,k}^j} e^{-\Psi_{0,k}}\,dm_{F_0}} \leqslant B_0\left(\frac{m_{F_0}\left(\left(\frac{1}{1+\eta}\right)\mathrm{C}_{0,k}^j\right)}{m_{F_0}(\mathrm{C}_{0,k}^j)}\right).$$

Hence, by the choice of η above, we have

$$\frac{m_{F_0}\left(\left(\frac{1}{1+\eta}\right)\mathrm{C}^j_{0,k}\right)}{m_{F_0}(\mathrm{C}^j_{0,k})} \geqslant \frac{1}{B_0\left(1+\frac{\varepsilon}{4}\right)}.$$

□

Notice that the above proof can be used to show that for the 1-cylinders considered above,

$$\frac{m_{F_0}\left(\mathrm{C}^j_{0,1}\setminus\left(\frac{1}{1+\eta}\right)\mathrm{C}^j_{0,1}\right)}{m_{F_0}(\mathrm{C}^j_{0,1})} \leqslant \frac{B_0}{1+\frac{\varepsilon}{4}}.$$

Proposition 4.5. *For all $\varepsilon>0$, $\lambda\in(0,1)$, $k_0\geqslant 1$ and sequences $(i_1,\ldots,i_{k_0})\in\mathbb{N}^{k_0}$ there exists $N_0\geqslant 1$ such that for all $n\geqslant N_0$, $1\leqslant k\leqslant k_0$ and $1\leqslant i\leqslant i_k$, we have*

$$\frac{1}{B_0^2(1+\varepsilon)} \leqslant \frac{m_{F_n}(\mathrm{C}^i_{0,k})}{e^{\Psi_{0,k}(x)}} \leqslant B_0^2(1+\varepsilon)$$

for all $x\in\lambda\mathrm{C}^j_{0,k}$.

Proof. The following claim is left to the reader.

Claim 3. *For all $\varepsilon>0$, $k_0\geqslant 1$ and sequences $(i_1,\ldots,i_{k_0})\in\mathbb{N}^{k_0}$ there exists $N_0\geqslant 1$ such that for all $n\geqslant N_0$, $1\leqslant k\leqslant k_0$ and $1\leqslant i\leqslant i_k$, $\mathrm{C}^i_{n,k}$ is matched. Moreover, for these cylinders, each set $F_n^{k-1}(\mathrm{C}^i_{n,k})$ is matched.*

We next make the following claim.

Claim 4. *For all $\varepsilon>0$, $k_0\geqslant 1$ and sequences $(i_1,\ldots,i_{k_0})\in\mathbb{N}^{k_0}$ there exist $\eta>0$ and $N_1\geqslant N_0$ such that for all $n\geqslant N_1$, $1\leqslant k\leqslant k_0$ and $1\leqslant i\leqslant i_k$,*

$$\frac{m_{F_n}((1+\eta)\mathrm{C}^i_{n,k})}{m_{F_n}(\mathrm{C}^i_{n,k})} \leqslant B_0\left(1+\frac{\varepsilon}{4}\right) \text{ and } \frac{m_{F_n}\left(\left(\frac{1}{1+\eta}\right)\mathrm{C}^i_{n,k}\right)}{m_{F_n}(\mathrm{C}^i_{n,k})} \geqslant \frac{1}{B_0\left(1+\frac{\varepsilon}{4}\right)}.$$

Proof. The proof of the claim is the same as for Lemma 4.4 except that we need to take F_n sufficiently close to F_0 so that the cylinders $\mathrm{C}^i_{n,k}$ considered in Lemma 4.4 have almost exactly the same properties as those $\mathrm{C}^i_{0,k}$ considered here. □

A simple consequence of these claims is that for all $\varepsilon>0$, $k_0\geqslant 1$ and sequences $(i_1,\ldots,i_{k_0})\in\mathbb{N}^{k_0}$ there exists $N_2\geqslant N_1$ such that for all $n\geqslant N_2$, $1\leqslant k\leqslant k_0$ and $1\leqslant i\leqslant i_k$, $\mathrm{C}^i_{n,k}\in\mathcal{P}_{n,k}$ is matched and

$$\frac{1}{B_0\left(1+\frac{\varepsilon}{4}\right)} \leqslant \frac{m_{F_n}(\mathrm{C}^i_{0,k})}{m_{F_n}(\mathrm{C}^i_{n,k})} \leqslant B_0\left(1+\frac{\varepsilon}{4}\right).$$

Here we choose $N_2\geqslant N_1$ so that $\mathrm{C}^j_{0,k}\subset(1+\eta)\mathrm{C}^j_{n,k}$ and $\mathrm{C}^j_{n,k}\subset(1+\eta)\mathrm{C}^j_{0,k}$ for all $\mathrm{C}^j_{n,k}$ as in the statement of the proposition.

The Gibbs property for m_{F_n}, which follows directly from conformality, means that $m_{F_n}(\mathrm{C}^i_{n,k}) = B_0^{\pm} e^{\Psi_{n,k}(x)}$ for all $x \in \mathrm{C}^i_{n,k}$. Now we can take N_2 so large that

$$\frac{1}{(1+\frac{\varepsilon}{4})} \leqslant e^{\Psi_{n,k}(x)-\Psi_{0,k}(x)} = e^{\Phi_{F_n,k}(x)-\Phi_{F_0,k}(x)+\left(P(\varphi_{f_n,t})-P(\varphi_{f_0,t})\right)\tau_{0,k}(x)} \leqslant 1+\frac{\varepsilon}{4}$$

for all $x \in \mathrm{C}^i_{n,k} \cap \mathrm{C}^i_{0,k}$ for the cylinders $\mathrm{C}^i_{n,k}$ under consideration. This follows since $\Phi_{F_n,k}(x) \to \Phi_{F_0,k}(x)$ as $n \to \infty$, and by Proposition 1.1, $P(\varphi_{f_n,t}) \to P(\varphi_{f_0,t})$ as $n \to \infty$. To complete the proof of the proposition, we possibly increase N_2 again to ensure that $\mathrm{C}^i_{n,k} \cap \mathrm{C}^i_{0,k} \subset \lambda \mathrm{C}^i_{0,k}$ for all the cylinders we consider. □

Combining Lemma 3.3 and Proposition 4.5 we have that μ_{F_∞} must have the Gibbs property with uniform constant. That is:

Corollary 4.6. *For all k and all $\mathrm{C}_{0,k} \in \mathcal{P}_{0,k}$,*

$$\frac{1}{H_{\mathcal{F}} B_0^2(1+\varepsilon)} \leqslant \frac{\mu_{F_\infty}(\mathrm{C}^i_{0,k})}{e^{\Psi_{0,k}(x)}} \leqslant H_{\mathcal{F}} B_0^2(1+\varepsilon),$$

for all $x \in \mathrm{C}^i_{0,k}$.

We will need the following lemma later.

Lemma 4.7. *For all $\varepsilon > 0$ and $i_0 \geqslant 1$ there exists $N \geqslant 1$ such that $n \geqslant N$ implies*

$$\mu_{F_n}\left(\bigcup_{i=0}^{i_0} \left(\mathrm{C}^i_{n,1} \triangle \mathrm{C}^i_{0,1}\right)\right) \leqslant \varepsilon.$$

Proof. Combining the arguments in the proof of Lemma 4.4, the paragraph following it and Claim 4 in the proof of Proposition 4.5 we have $\eta > 0$, $i_0 \geqslant 1$ and $N' \geqslant 1$ such that for $n \geqslant N'$,

$$m_{F_n}\left((1+\eta)\mathrm{C}^i_{n,k} \setminus \mathrm{C}^i_{n,k}\right),\ m_{F_n}\left(\mathrm{C}^i_{n,k} \setminus \frac{\mathrm{C}^i_{n,k}}{(1+\eta)}\right) < \frac{\varepsilon}{i_0 H_{\mathcal{F}}}$$

for all $1 \leqslant i \leqslant i_0$. Recall that $H_{\mathcal{F}}$ is the constant from Lemma 3.3. Moreover, there exists $N \geqslant N'$ such that $n \geqslant N$ implies

$$\mathrm{C}^i_{n,1} \triangle \mathrm{C}^i_{0,1} \subset (1+\eta)\mathrm{C}^i_{n,k} \setminus \frac{\mathrm{C}^i_{n,k}}{(1+\eta)}$$

for all $1 \leqslant i \leqslant i_0$. Therefore, $n \geqslant N$ implies

$$m_{F_n}\left(\bigcup_{i=0}^{i_0} \left(\mathrm{C}^i_{n,1} \triangle \mathrm{C}^i_{0,1}\right)\right) \leqslant \frac{\varepsilon}{H_{\mathcal{F}}}.$$

The lemma follows from Lemma 3.3, substituting $\mu_{F_n} (= \rho_{F_n} m_{F_n})$ for m_{F_n} in the above equation. □

We finish this section by proving Proposition 1.1, which was essential in the proof of Proposition 4.5.

Proof of Proposition 1.1. Observe that for $t = 1$ there is nothing to prove since $P(\varphi_{f,1}) = 0$ for all $f \in \mathcal{F}$.

Let $\varepsilon > 0$. We fix $t < 1$ as in the statement of the proposition, since the proof for $t > 1$ (which we need only consider when (1) holds) follows similarly. For ease of notation, we let $P_n := P(\varphi_{f_n,t})$. For $S \in \mathbb{R}$, we define $\psi_n^S := \varphi_{f_n,t} - S$. Likewise, for $k \geqslant 1$ the corresponding induced potentials are $\Psi_{n,k}^S := \Phi_{n,k} - S\tau_{n,k}$.

We choose $\mathrm{C}_{0,1}^i$ and $n_1 \geqslant 1$ so that for all $n \geqslant n_1$, this cylinder is matched. Recall that we can write

$$P_G\left(\Psi_n^{P_n}\right) = P_G\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right) = \lim_{k\to\infty} \frac{1}{k} \log Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right) = 0,$$

where $Z_k(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i) = \sum_{x\in \mathrm{C}_{n,1}^i,\ F_n^k(x)=x} e^{\Psi_{n,k}^{P_n}(x)}$. The idea of this proof is to use the fact that $S = P_n$ is the unique value so that $P_G(\Psi_{n,1}^S, \mathrm{C}_{n,1}^i) = 0$. We show that since $\frac{1}{k}\log Z_k(\Psi_0^{P_n}, \mathrm{C}_{0,1}^i)$ and $\frac{1}{k}\log Z_k(\Psi_n^{P_n}, \mathrm{C}_{n,1}^i)$ are close to each other for all large n, with the latter value converging to 0 as $k \to \infty$, then P_0 and P_n must also be close.

We first show that the convergence of $\frac{1}{k}\log Z_k(\Psi_n^{P_n}, \mathrm{C}_{n,1}^i)$ to 0 as $k \to \infty$ is essentially uniform in n. The Gibbs property together with Lemma 3.3 imply that for a cylinder $\mathrm{C}_{n,k} \in \mathcal{P}_{n,k}$ we have $\mu_{\Psi_n}^{P_n}(\mathrm{C}_{n,k}) = \mathcal{H}_{\mathcal{F}}^{\pm} e^{\Psi_{n,k}(x)}$ for any $x \in \mathrm{C}_{n,k}$, Since each k-cylinder contains a unique k-periodic point, it follows that

$$Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right) = H_{\mathcal{F}}^{\pm} \sum_{\mathrm{C}_{n,k}\in\mathcal{P}_{n,k}, \mathrm{C}_{n,k}\subset \mathrm{C}_{n,1}} \mu_{\Psi_n^{P_n}}(\mathrm{C}_{n,k}) = H_{\mathcal{F}}^{\pm} \mu_{\Psi_n^{P_n}}(\mathrm{C}_{n,1}^i).$$

We now show that $\mu_{\Psi_n^{P_n}}(\mathrm{C}_{n,1})$ is uniformly bounded above and below for all large n. Again using the Gibbs property we have

$$\frac{1}{H_{\mathcal{F}}} \leqslant \frac{\mu_{\Psi_n^{P_n}}(\mathrm{C}_{n,1}^i)}{e^{\Phi_n(x) - P_n\tau_n(x)}} \leqslant H_{\mathcal{F}}$$

for any $x \in \mathrm{C}_{n,1}^i$. Since there exists a uniform $K_{\mathcal{F}} \in \mathbb{R}$ so that for all $n \geqslant 0$, $P_n \in [0, K_{\mathcal{F}}]$; $\mathrm{C}_{n,1}^i$ is matched (recall the definition on 16), for all $n \geqslant n_1$; and Φ_n converges to Φ_0 on $\mathrm{C}_{n,1}^i \cap \mathrm{C}_{0,1}^i$, there exists $n_2 \geqslant n_1$ so that $n = 0$ or $n \geqslant n_2$ implies

$$\frac{e^{-K_{\mathcal{F}}\tau_0(x)}}{H_{\mathcal{F}}(1+\varepsilon)} \leqslant \frac{\mu_{\Psi_n^{P_n}}(\mathrm{C}_{n,1}^i)}{e^{\Phi_0(x)}} \leqslant H_{\mathcal{F}}(1+\varepsilon)$$

for any $x \in \mathrm{C}_{0,1}^i$. Combining the above computations we get, for $n = 0$ or $n \geqslant n_2$,

$$\frac{e^{\Phi_0(x) - K_{\mathcal{F}}\tau_0(x)}}{H_{\mathcal{F}}^2(1+\varepsilon)} \leqslant Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right) \leqslant H_{\mathcal{F}}^2(1+\varepsilon)e^{\Phi_0(x)} \tag{7}$$

for any $x \in \mathrm{C}_{0,1}^i$.

Therefore for any $\delta > 0$ there exists a uniform $k = k(\delta) \geqslant 1$ so that

$$\left|\frac{1}{k}\log Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right)\right| < \delta$$

for $n = 0$ and all $n \geqslant n_2$. We fix $\delta = \frac{\varepsilon}{5}$.

In order to be able to prove this proposition using only a finite amount of information, we define

$$Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i, N\right) := \sum_{x\in \mathrm{C}_{n,1}^i,\ F^k(x)=x,\ \tau_{n,k}(x)\leqslant N} e^{\Psi_{n,k}^{P_n}(x)}.$$

By the Gibbs property, the quantity $Z_k(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i) - Z_k(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i, N)$ is, up to a constant, the measure of the 'tail set for F_n^k', i.e. $\mu_{\Psi_n^{P_n}}\{\tau_{n,k} > N\}$. It is easy to see that estimating this last quantity can be reduced to estimating $\mu_{\Psi_n^{P_n}}\{\tau_n > N\}$ (see for example [ACF, Lemma 9.7]). Therefore the following claim follows for the same reasons as in Lemma 4.1. (Recall here that we are not dealing with the case $t = 1$, so even if we only assume that (2) holds on $\mathcal{F}$, since we are considering $t < 1$, we have exponential tails.) In contrast to the notation in the rest of this proof, here, for clarity, we emphasise the role of $\mathcal{F}$ and t.

Claim 5. *There exist $C_{\mathcal{F},t}, \alpha_{\mathcal{F},t} > 0$ so that for $n = 0$ or $n \geqslant n_2$, and for all $k \geqslant 0$,*

$$Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i\right) - Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i, N\right) \leqslant C_{\mathcal{F},t} e^{-\alpha_{\mathcal{F},t} N}.$$

Adding (7) and the claim together, we can fix $N = N(\delta, n_1) \geqslant 1$ so that for all $n \geqslant n_2$,

$$e^{-3\delta k} \leqslant \frac{Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i, N\right)}{Z_k\left(\Psi_{0,1}^{P_0}, \mathrm{C}_{0,1}^i, N\right)} \leqslant e^{3\delta k}. \tag{8}$$

Now since $\Psi_{n,k}^S$ on the set $\{\tau_{0,k} \leqslant N\}$ essentially converges to $\Psi_{0,k}^S$ as $n \to \infty$, there exists $n_3 \geqslant n_2$ so that $n \geqslant n_3$ implies $e^{-k\delta} \leqslant \frac{Z_k\left(\Psi_{0,1}^{P_n}, \mathrm{C}_{0,1}^i, N\right)}{Z_k\left(\Psi_{n,1}^{P_n}, \mathrm{C}_{n,1}^i, N\right)} \leqslant e^{k\delta}$, and hence

$$e^{-4\delta k} \leqslant \frac{Z_k\left(\Psi_{0,1}^{P_n}, \mathrm{C}_{0,1}^i, N\right)}{Z_k\left(\Psi_{0,1}^{P_0}, \mathrm{C}_{0,1}^i, N\right)} \leqslant e^{4\delta k}.$$

But $\tau_{0,k} \geqslant k$, so by the definition of $Z_k\left(\Psi_{0,1}^S, \mathrm{C}_{0,1}^i, N\right)$, the above inequalities imply that $|P_n - P_0|$ must be less than or equal to $4\delta < \varepsilon$ as required. $\square$

5. Invariance of the weak* limit

We may assume, as in the beginning of Section 4, that μ_{F_∞} is the weak* limit of the sequence $(\mu_{F_n})_n$. In the previous section we saw that μ_{F_∞} is Gibbs. The purpose of this section is to show that μ_{F_∞} is F_0-invariant. Before that, we prove the following technical lemma that will be useful in the remaining arguments.

Lemma 5.1. *For all $i \in \mathbb{N}$ and every continuous $g : C_{0,1}^i \to \mathbb{R}$ we have*

$$\int g.\mathbf{1}_{C_{0,1}^i}\, d\mu_{F_n} \to \int g.\mathbf{1}_{C_{0,1}^i}\, d\mu_{F_\infty}.$$

Proof. We can extend g continuously to $\partial C^i_{0,1}$, and for every $x \in I \setminus \overline{C^i_{0,1}}$, define $\mathrm{b}^i(x)$ as the point of $\partial C^i_{0,1}$ closest to x.

Observing that $g = g^+ - g^-$, where $g^+(x) = \max\{0, g(x)\} \geqslant 0$ and $g^-(x) = \max\{0, -g(x)\} \geqslant 0$, we may assume without loss of generality that $g \geqslant 0$. Also, since, by Corollary 4.6, μ_{F_∞} is a Gibbs measure, we have $\mu_{F_\infty}(\partial C^i_{0,1}) = 0$, which implies that $\int_{\overline{C^i_{0,1}} \setminus \partial C^i_{0,1}} g \, d\mu_{F_\infty} = \int_{\overline{C^i_{0,1}}} g d\mu_{F_\infty} = \int_{C^i_{0,1}} g \, d\mu_{F_\infty}$.

Let $U_k = \{x \in I : \mathrm{dist}(x, \overline{C^i_{0,1}}) < 1/k\}$. Clearly U_k is an open neighbourhood of $\overline{C^i_{0,1}}$ and by the regularity of μ_{F_∞} it follows that $\mu_{F_\infty}(U_k \setminus \overline{C^i_{0,1}}) = \epsilon(k) \to 0$ as $k \to \infty$. Define $h : I \to \mathbb{R}$ as

$$h(x) = \begin{cases} 0 & \text{if } x \notin U_k \\ g(\mathrm{b}^i(x)) \frac{d(x, I \setminus U_k)}{d(x, I\setminus U_k) + d(x, \overline{C^i_{0,1}})} & \text{if } x \in U_k \setminus \overline{C^i_{0,1}} \\ g(x) & \text{if } x \in \overline{C^i_{0,1}} \end{cases}.$$

Notice that h is continuous and, for every $x \in I$, we have $g(x)\mathbf{1}_{C^i_{0,1}}(x) \leqslant h(x) \leqslant \max_{x \in \overline{C^i_{0,1}}} g(x)$ and $h(x) - g(x)\mathbf{1}_{C^i_{0,1}}(x) > 0$ only if $x \in U_k \setminus \overline{C^i_{0,1}}$. Consequently, using the weak* convergence of μ_{F_n} to μ_{F_∞}, it follows

$$\int g \mathbf{1}_{C^i_{0,1}} d\mu_{F_n} \leqslant \int h \, d\mu_{F_n} \xrightarrow[n\to\infty]{} \int h \, d\mu_{F_\infty} \leqslant \int g \mathbf{1}_{\overline{C^i_{0,1}}} \, d\mu_{F_\infty} + \epsilon(k) \max_{x \in \overline{C^i_{0,1}}} g(x).$$

Letting $k \to \infty$ we get $\int g \mathbf{1}_{C^i_{0,1}} d\mu_{F_n} \leqslant \int g \mathbf{1}_{C^i_{0,1}} d\mu_{F_\infty}$. The opposite inequality follows similarly. □

Lemma 5.2. *μ_{F_∞} is F_0-invariant.*

Proof. The F_0-invariance of μ_{F_∞} is equivalent to

$$\int \varphi \circ F_0 \, d\mu_{F_\infty} = \int \varphi \, d\mu_{F_\infty}$$

for every continuous $\varphi \colon I \to \mathbb{R}$. Given any $\varphi \colon I \to \mathbb{R}$ continuous we have by hypothesis

$$\int \varphi \, d\mu_{F_n} \to \int \varphi \, d\mu_{F_\infty} \quad \text{as} \quad n \to \infty.$$

On the other hand, since μ_{F_n} is an F_n-invariant probability measure, we have

$$\int \varphi \, d\mu_{F_n} = \int (\varphi \circ F_n) \, d\mu_{F_n} \quad \text{for every } n \geqslant 0.$$

So, it suffices to prove that

$$\int (\varphi \circ F_n) \, d\mu_{F_n} \to \int (\varphi \circ F_0) \, d\mu_{F_\infty} \quad \text{as} \quad n \to \infty. \tag{9}$$

We have

$$\left| \int (\varphi \circ F_n) \, d\mu_{F_n} - \int (\varphi \circ F_0) \, d\mu_{F_\infty} \right| \leqslant$$
$$\left| \int (\varphi \circ F_n) \, d\mu_{F_n} - \int (\varphi \circ F_0) \, d\mu_{F_n} \right| + \left| \int (\varphi \circ F_0) \, d\mu_{F_n} - \int (\varphi \circ F_0) \, d\mu_{F_\infty} \right|.$$

Observing that $\varphi \circ F_0$ is continuous on each $C^i_{0,1}$, we easily deduce from Lemma 5.1 and Lemma 4.3 that the second term in the sum above is close to zero for large n.

The only thing we are left to prove is that the first term in the sum above converges to 0 when n tends to ∞. That term is bounded above by

$$\int |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n}. \tag{10}$$

Take any $\varepsilon > 0$. Using Lemma 4.1, take $N \geqslant 1$ such that

$$\sum_{\tau_n^i > N} \mu_{F_n}(C^i_{n,1}) < \varepsilon.$$

We write the integral in (10) as

$$\sum_{\tau_n^i > N} \int_{C^i_{n,1}} |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n} + \sum_{\tau_n^i \leqslant N} \int_{C^i_{n,1}} |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n}. \tag{11}$$

The first sum in (11) is bounded by $2\varepsilon\|\varphi\|_\infty$. Let us now estimate the second sum in (11).

Using Lemma 4.3, we take n_1 sufficiently large so that for all $n \geqslant n_1$ and every cylinder $C^i_{n,1}$ with $\tau_n^i \leqslant N$ there is a matching cylinder $C^i_{0,1}$ with $\tau_n^i = \tau_0^i$. Moreover, using Lemma 4.7, we may assume that n_1 is large enough so that $n \geqslant n_1$ implies

$$\sum_{\tau_n^i \leqslant N} \mu_{F_n}(C^i_{n,1} \triangle C^i_{0,1}) < \varepsilon.$$

For every i such that $\tau_n^i \leqslant N$ we have

$$\begin{aligned}\int_{C^i_{n,1}} |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n} \leqslant & \int_{C^i_{n,1} \cap C^i_{0,1}} |\varphi \circ f_n^{\tau_0^i} - \varphi \circ f_0^{\tau_0^i}| \, d\mu_{F_n} \\ & + \int_{C^i_{n,1} \setminus C^i_{0,1}} |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n}.\end{aligned}$$

Since $f_n \to f_0$ in the C^k topology, there is $n_2 \in \mathbb{N}$ such that for $n \geqslant n_2$

$$\sum_{\tau_n^i \leqslant N} \int_{C^i_{n,1} \cap C^i_{0,1}} |\varphi \circ f_n^{\tau_n^i} - \varphi \circ f_0^{\tau_n^i}| \, d\mu_{F_n} < \varepsilon.$$

On the other hand, for $n \geqslant n_1$

$$\sum_{\tau_n^i \leqslant N} \int_{C^i_{n,1} \triangle C^i_{0,1}} |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n} \leqslant 2\varepsilon\|\varphi\|_\infty.$$

Thus we have for $n \geqslant \max\{n_1, n_2\}$

$$\int |\varphi \circ F_n - \varphi \circ F_0| \, d\mu_{F_n} \leqslant \varepsilon(4\|\varphi\|_\infty + 1).$$

This proves the result since $\varepsilon > 0$ was arbitrary. □

Since μ_{F_∞} is an invariant Gibbs measure, uniqueness of such measures, [MU, Theorem 3.2], implies $\mu_{F_\infty} \equiv \mu_{F_0}$.

Remark 5.3. *Observe that the whole sequence μ_{F_n} converges in the weak* topology to μ_{F_0}. This is because any subsequence $\left(\mu_{F_{n_i}}\right)_i$ admits a convergent subsequence $\left(\mu_{F_{n_{i_j}}}\right)_j$, whose weak* limit, μ_{F_∞}, is Gibbs and F_0-invariant, by Corollary 4.6 and Lemma 5.2. Hence, by uniqueness, $\mu_{F_{n_{i_j}}} \to \mu_{F_0}$, in the weak* topology, which clearly implies the statement.*

6. Continuous variation of equilibrium states

So far, we managed to prove that if $\|f_n - f_0\|_{C^2} \to 0$, then the induced Gibbs measures converge in the weak* topology, ie, $\mu_{F_n} \to \mu_{F_0}$. Similarly to (4), we let $\mu_f^*/\int \tau d\mu_F$ be the projection of the measure μ_F, i.e.

$$\mu_f^* = \sum_{i=1}^{\infty} \sum_{k=0}^{\tau_i - 1} f_*^k \left(\mu_F | C_1^i\right) \tag{12}$$

By Lemma 4.1, the total mass of this measure $\mu_{f_n}^*(I)$ is uniformly bounded in n. Observe that, for some fixed $t \in U_{\mathcal{F}}$, the unique equilibrium state of f_n for the potential $-t \log |Df_n|$ is such that $\mu_n = \mu_{f_n}^*/\mu_{f_n}^*(I)$, for every $n \geqslant 0$. Consequently, the proof of Theorem A will be complete once we prove:

Proposition 6.1. *For every continuous $g : I \to I$,*

$$\int g \; d\mu_{f_n}^* \xrightarrow[n\to\infty]{} \int g \; d\mu_{f_0}^*.$$

Proof. First observe that as I is compact, g is uniformly continuous and $\|g\|_\infty < \infty$.

Let $\varepsilon > 0$ be given. We look for $n_0 \in \mathbb{N}$ sufficiently large so that for every $n > n_0$

$$\left| \int g \; d\mu_{f_n}^* - \int g \; d\mu_{f_0}^* \right| < \varepsilon$$

Recalling (12) we may write for any integer N

$$\mu_{f_n}^* = \sum_{\tau_n^i \leqslant N} \sum_{k=0}^{\tau_n^i - 1} (f_n^k)_* (\mu_{F_n} | C_{n,1}^i) + \eta_{f_n} \text{ and } \mu_{f_0}^* = \sum_{\tau_0^i \leqslant N} \sum_{k=0}^{\tau_0^i - 1} (f_0^k)_* (\mu_{F_0} | C_{0,1}^i) + \eta_{f_0}$$

where $\eta_{f_n} = \sum_{\tau_n^i > N} \sum_{k=0}^{\tau_n^i - 1} (f_n^k)_* (\mu_{F_n} | C_{n,1}^i)$ and $\eta_{f_0} = \sum_{\tau_0^i > N} \sum_{k=0}^{\tau_0^i - 1} (f_0^k)_* (\mu_{F_0} | C_{0,1}^i)$. Using Lemma 4.1 we pick N large enough so that $n \geqslant N$ implies

$$\eta_{f_n}(I) + \eta_{f_0}(I) < \varepsilon/2.$$

Using Lemma 4.3, we take n_1 sufficiently large so that for all $n \geqslant n_1$ and every cylinder $C_{n,1}^i$ with $\tau_n^i \leqslant N$ there is a matching cylinder $C_{0,1}^i$ with $\tau_n^i = \tau_0^i$. Let S_N denote the number of 1-cylinders such that $\tau_n^i \leqslant N$. To complete the proof of the proposition, for every i such that $\tau_n^i \leqslant N$ and $k < \tau_n^i$, we must find a sufficiently large n_2 so that for every $n \geqslant n_2$

$$E := \left| \int (g \circ f_n^k) \mathbf{1}_{C_{n,1}^i} \; d\mu_{F_n} - \int (g \circ f_0^k) \mathbf{1}_{C_{0,1}^i} \; d\mu_{F_0} \right| < \frac{\varepsilon}{2 S_N}.$$

We split E into E_1, E_2 and E_3 presented in respective order:

$$\begin{aligned} E \leqslant & \left| \int \left[(g \circ f_n^k) - (g \circ f_0^k) \right] \mathbf{1}_{C^i_{n,1}} d\mu_{F_n} \right| \\ & + \left| \int (g \circ f_0^k) \left[\mathbf{1}_{C^i_{n,1}} - \mathbf{1}_{C^i_{0,1}} \right] d\mu_{F_n} \right| \\ & + \left| \int (g \circ f_0^k) \mathbf{1}_{C^i_{0,1}} d\mu_{F_n} - \int (g \circ f_0^k) \mathbf{1}_{C^i_{0,1}} d\mu_{F_0} \right| . \end{aligned}$$

Since

$$E_1 \leqslant \int \left| (g \circ f_n^k) - (g \circ f_0^k) \right| \, d\mu_{F_n},$$

we choose n_2 large enough so that for every $n > n_2$ we have $\left| (g \circ f_n^k) - (g \circ f_0^k) \right| \leqslant \frac{\varepsilon}{6S_N}$ in order to obtain $E_1 \leqslant \frac{\varepsilon}{6S_N}$.

Now,

$$E_2 \leqslant \|g\|_\infty \mu_{F_n}(C^i_{n,1} \triangle C^i_{0,1}).$$

Using Lemma 4.7, we take n_2 large enough so that for all $n > n_2$ we have $E_2 \leqslant \frac{\varepsilon}{6S_N}$.

Regarding the last term, Lemma 5.1 allows us to conclude that if n_2 is sufficiently large then for all $n > n_2$ we have $E_3 \leqslant \frac{\varepsilon}{6S_N}$. □

Proof of Theorem B. Alves and Viana, in [AV], give some abstract conditions for statistical stability of physical measures in the strong sense, that is, convergence of densities in the sense of (3). Essentially, they consider a family $\mathcal{U}$ of C^k ($k \geq 2$) maps admitting an inducing scheme. Their main result, [AV, Therorem A], asserts that if some uniformity conditions, which they denote by U_1, U_2 and U_3, hold within the family, then one gets strong statistical stability. Condition U_1 requires that cylinders with finite inducing times are arbitrarily close with respect to the reference measure, just as we have shown for our inducing schemes in the proof of Lemma 4.7. Condition U_2 requires uniformity on the decay of the tail of the inducing times, which is covered, in our case, by Lemma 4.1. Condition U_3 demands that the constants involved on the estimates of the induced map (such as bounded distortion, derivative growth, backward contraction, etc.) can be chosen uniformly in a neighbourhood of each map $f \in \mathcal{U}$. This also holds in the present situation as it has been discussed during the proof of Lemma 4.1.

Consequently, our inducing schemes and their properties put us trivially in the setting of Alves and Viana, meaning that both $\mathcal{F}_e(r, \ell, C, \alpha)$ and $\mathcal{F}_p(r, \ell, C, \beta)$ meet all the requirements of the family $\mathcal{U}$ in [AV, Therorem A], from which we conclude that

$$\mathcal{F} \ni f \mapsto \frac{d\mu_f}{dm}$$

is continuous as in (3), where $\mathcal{F}$ stands for either $\mathcal{F}_e(r, \ell, C, \alpha)$ or $\mathcal{F}_p(r, \ell, C, \beta)$ and m denotes Lebesgue measure. □

Remark 6.2. *Note that the theory presented here extends to Manneville-Pomeau maps $f : x \mapsto x + x^{1+\alpha}$ (mod 1) for $\alpha \in (0, 1)$. Given such a map, and a potential $\varphi_t := -t \log |Df|$, it is straightforward to prove an equivalent of* [BT2, Theorem 1],

yielding an equilibrium state μ_t for $t \in [\delta, 1]$ for some $\delta < 0$. One main difference in proving statistical stability for these measures is that in the proofs of Proposition 4.5 and Lemma 4.7 for example, to estimate the measure of sets $\mathrm{C}^i_{0,k} \triangle \mathrm{C}^i_{n,k}$ we can no longer assume that no two cylinders for the inducing schemes are adjacent. Above, this property enabled us to estimate $\mathrm{C}^i_{0,k} \triangle \mathrm{C}^i_{n,k}$ using the measure of 1-cylinders. However, when, as in the Manneville-Pomeau case, we do not have this property, we can use the measure of k-cylinders to give us the required estimates instead.

In broad terms, the theory presented here will also go through in more general families of maps, for example to simple generalisations of Manneville-Pomeau maps. As we have seen, the important ingredients are that the inducing schemes for the families can be chosen so that for nearby maps, the inducing schemes are 'close'; that the thermodynamic formalism in Section 3 goes through for the inducing schemes; and that there are uniform bounds for the decay of the tail sets (as in Lemma 4.1) for all the inducing schemes.

References

[ACF] J. F. Alves, M. P. Carvalho, J. M. Freitas, *Statistical Stability for Hénon maps of the Benedicks-Carleson type,* Preprint, arXiv:math/0610602.

[AV] J.F. Alves, M. Viana, *Statistical stability for robust classes of maps with non-uniform expansion,* Ergodic Theory Dynam. Systems **22** (2002) 1–32.

[Ar] V. Araújo, *Semicontinuity of entropy, existence of equilibrium states and continuity of physical measures,* Discrete Contin. Dyn. Syst. **17** (2007) 371–386.

[Ba1] V. Baladi, *On the susceptibility function of piecewise expanding interval maps,* Comm. Math. Phys. **275** (2007) 839–859

[Ba2] V. Baladi, *Linear response despite critical points,* Nonlinearity **21** (2008) T81–T90.

[B] H. Bruin, *Induced maps, Markov extensions and invariant measures in one–dimensional dynamics,* Commun. Math. Phys. **168** (1995) 571–580.

[BK] H. Bruin, G. Keller, *Equilibrium states for S–unimodal maps,* Ergodic Theory Dynam. Systems **18** (1998) 765–789.

[BLS] H. Bruin, S. Luzzatto, S. van Strien, *Decay of correlations in one–dimensional dynamics*, Ann. Sci. École Norm. Sup. **36** (2003) 621–646.

[BT1] H. Bruin, M. Todd, *Markov extensions and lifting measures for complex polynomials,* Ergodic Theory Dynam. Systems **27** (2007) 743–768.

[BT2] H. Bruin, M. Todd, *Equilibrium states for interval maps: the potential* $-t \log |Df|$, Preprint, arXiv:0704.2199.

[BT3] H. Bruin, M. Todd, *Equilibrium states for potentials with* $\sup \varphi - \inf \varphi < h_{top}(f)$, Comm. Math. Phys. **283** (2008) 579–611.

[F] J. Freitas, *Continuity of SRB measure and entropy for Benedicks-Carleson quadratic maps,* Nonlinearity **18** (2005) 831–854.

[H] F. Hofbauer, *The topological entropy of a transformation* $x \mapsto ax(1-x)$, Monath. Math. **90** (1980) 117–141.

[HK] F. Hofbauer, G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations,* Math. Z. **180** (1982) 119–140.

[HR] F. Hofbauer, P. Raith, *Topologically transitive subsets of piecewise monotonic maps which contain no periodic points,* Monatsh. Math. **107** (1989) 217–239.

[K] G. Keller, *Lifting measures to Markov extensions,* Monatsh. Math. **108** (1989) 183–200.

[KN] G. Keller, T. Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps,* Comm. Math. Phys. 149 (1992) 31–69.

[L] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval,* Ergodic Theory Dynam. Systems **1** (1981) 77–93.

[MU] D. Mauldin, M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet,* Israel J. Math. **125** (2001) 93–130.

[MeS] W. de Melo, S. van Strien, *One dimensional dynamics,* Ergebnisse Series **25**, Springer–Verlag, 1993.

[MSz] M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings,* Studia Math. **67** (1980) 45–63.

[P] Y. Pesin, *Dimension theory in dynamical systems,* University of Chicago Press, Chicago, IL, 1997.

[PSe] Y. Pesin, S. Senti, *Equilibrium measures for maps with inducing schemes,* J. Mod. Dyn. **2** (2008) 397–427.

[PW] Y. Pesin, H. Weiss, *The multifractal analysis of Birkhoff averages and large deviations,* Global analysis of dynamical systems, 419–431, Inst. Phys., Bristol, 2001.

[Ru] D. Ruelle, *Differentiation of SRB states,* Comm. Math. Phys. **187** (1997) 227–241.

[RS] M. Rychlik, E. Sorets, *Regularity and other properties of absolutely continuous invariant measures for the quadratic family*, Comm. Math. Phys. **150** (1992) 217–236.

[Sa1] O. Sarig, *Thermodynamic formalism for countable Markov shifts,* Ergodic Theory Dynam. Systems **19** (1999) 1565–1593.

[Sa2] O. Sarig, *Thermodynamic formalism for Markov shifts,* PhD. thesis, Tel–Aviv, 2000.

[Sa3] O. Sarig, *Phase transitions for countable Markov shifts,* Commun. Math. Phys. **217** (2001) 555-577.

[Sa4] O. Sarig, *Existence of Gibbs measures for countable Markov shifts,* Proc. Amer. Math. Soc. **131** (2003) 1751–1758.

[Th] H. Thunberg, *Unfolding of chaotic unimodal maps and the parameter dependence of natural measures*, Nonlinearity **14** (2001) 323–337.

[T] M. Todd, *Multifractal analysis for multimodal maps,* Preprint, arXiv:0809.1074.

[Ts] M. Tsujii, *On continuity of Bowen-Ruelle-Sinai measures in families of one dimensional maps*, Comm. Math. Phys. **177** (1996) 1–11.

[W] P. Walters, *An introduction to ergodic theory,* Springer-Verlag, 1982.

Centro de Matemática da Universidade do Porto
Rua do Campo Alegre 687
4169-007 Porto
Portugal
jmfreita@fc.up.pt, http://www.fc.up.pt/pessoas/jmfreita/
mtodd@fc.up.pt, http://www.fc.up.pt/pessoas/mtodd/